

УНИВЕРЗИТЕТ „ГОЦЕ ДЕЛЧЕВ“ –ШТИП

ФАКУЛТЕТ ЗА ИНФОРМАТИКА



Докторска дисертација

**СИСТЕМ ЗА АМБИЕНТАЛНО ПОМОГНАТО ЖИВЕЕЊЕ И КОНТИНУИРАНО
СЛЕДЕЊЕ НА БИОФИЗИОЛОШКИ ПАРАМЕТРИ**

Александра Стојанова

Штип, 2020

Комисија за оценка на одбраната

Интерен ментор: Проф. д-р. Цвета Мартиновска Банде

Редовен професор на Факултет за информатика,
Универзитет „Гоце Делчев“

Екстерен ментор: Проф. д-р. Андреа Кулаков

Редовен професор на Факултет за информатички науки и
компјутерско инженерство, Универзитет „Св. Кирил и Методиј“

Членови на комисијата за оценка и одбрана

Претседател: Проф. д-р. Сашо Коцески

Редовен професор на Факултет за информатика,
Универзитет „Гоце Делчев“

Член: Проф. д-р. Цвета Мартиновска Банде

Редовен професор на Факултет за информатика,
Универзитет „Гоце Делчев“

Член: Проф. д-р. Андреа Кулаков

Редовен професор на Факултет за информатички науки и
компјутерско инженерство, Универзитет „Св. Кирил и Методиј“

Член : Проф. д-р. Владо Гичев

Редовен професор на Факултет за информатика,
Универзитет „Гоце Делчев“

Член: Проф. д-р Милка Здравковска

Редовен професор на Факултет за медицински науки,
Универзитет „Гоце Делчев“

Научно поле: 110 Информатика

212 Компјутерска техника и информатика

Научна област: 11000 Информатика

11011 Друго

21205 Вештачка интелигенција и системи

21206 Процесирање на податоци

Датум на одбрана 16.09.2020

Датум на промоција _____

Листа на трудови поврзани со темата

- [1] Stojanova, Aleksandra and Koceski, Saso and Koceska, Natasa (2018) Continuous Blood Pressure Monitoring as a Basis for Ambient Assisted Living (AAL) – Review of Methodologies and Devices. Journal of Medical Systems. ISSN 0148-5598
- [2] Stojanova, Aleksandra and Kocaleva, Mirjana and Luledjieva, Marija and Koceski, Saso (2019) High level activity recognition using android smart phone sensors - Review. Balkan Journal of Applied Mathematics and Informatics, 2 (2). pp. 27-36. ISSN 2545-4803
- [3] Stojanova, Aleksandra and Zlatanovska, Biljana and Kocaleva, Mirjana and Gicev, Vlado (2015) Obtaining functions from Fourier series with Matlab. A journal for information technology, education development and teaching methods of technical and natural sciences, 5 (1). pp. 1-10. ISSN 2217-7949
- [4] Kocaleva, Mirjana and Zlatanovska, Biljana and Stojanova, Aleksandra and Stojkovic, Natasa and Gicev, Vlado (2017) Wave equation with Dirichlet boundary conditions. In: ITRO 2017, 22 June 2017, Zrenjanin, Serbia.
- [5] Kocaleva, Mirjana and Stojanova, Aleksandra and Koceska, Natasa (2017) Review: Using physiological parameters for evaluating User Experience. In: ITRO 2017, 22 June 2017, Zrenjanin, Serbia.

Содржина

Листа на слики	vi
List of Figures.....	ix
Листа на табели.....	xi
List of Tables.....	xii
Резиме.....	xiii
Abstract.....	xv
1. Вовед.....	1
1.1 Биофизиолошки параметри	1
1.1.1 Електрокардиограм	3
1.1.2 Срцево отчукување (heartrate-HR).....	4
1.1.3 Крвен притисок	5
1.1.4 Фреквенција на дишење (RespirationRate -RR)	6
1.1.5 Кислородна заситеност (Blood Oxygen Saturation - SpO2)	7
1.1.6 Шеќер во крвта	8
1.1.7 Потење на кожата	9
1.1.8 Капнографија (Capnography)	10
1.1.9 Телесна температура.....	11
1.2 Континуирано следење на крвниот притисок.....	11
1.2.1 Можни методи за мерење на крвниот притисок.....	13
1.2.2 Методологии за континуирано мерење на крвниот притисок	16
1.2.2.1 Естимација на BP со користење на сигналите ЕКГ и PPG.....	20
1.2.2.2 Естимација на BP со користење само на PPG сигнали	21
1.2.2.3 Естимација на BP со користење само на ЕКГ сигнали	22
1.2.3 Уреди за континуирано мерење на крвниот притисок	22
1.3 Цели на истражувањето	23
1.4 Структурираност на трудот	25

2. Пресметување во магла (FOG computing).....	27
2.1. Карактеристики на пресметувањето во магла.....	28
2.2. Прашања и проблеми поврзани со пресметувањето во магла	32
2.2.1. Безбедност и приватност.....	32
2.2.2. Топологија на мрежата Fog и свесност на јазлите за локацијата.....	34
2.2.3. Управување со ресурси	34
2.2.4. Интероперабилност	35
2.2.5. Други проблеми	35
2.3. Пресметка во Облак наспроти пресметка во Магла	35
2.4. Поврзани парадигми и технологии	42
2.4.1. Mobile cloud computing (MCC)	42
2.4.2. Mobile edge computing (MEC).....	44
2.4.3. Edge computing	45
2.4.4. Dew computing (DC)	46
2.4.5. Fog-Dew computing	47
2.5. Fog Computing наспроти Edge computing	48
2.6. Архитектура на пресметувањето во магла	51
2.6.1. Високо ниво на Fog Computing архитектурата	51
2.6.2. Различни предложени архитектури за пресметката во магла	53
2.6.3. Компоненти на Fog Computing архитектура	53
2.6.3.1 Физичко ниво.....	55
2.6.3.2 Магла уреди, сервер и порта ниво	55
2.6.3.3 Ниво за набљудување (Monitoring layer)	56
2.6.3.4 Пред и по процесирачко ниво (Pre and post-processing layer)	57
2.6.3.5 Ниво на складирање (Storage layer)	58
2.6.3.6 Ниво за управување со ресурси (Resource management layer)	58
2.6.3.7 Безбедносно ниво (Security layer).....	59
2.6.3.8 Апликациско ниво (Application layer).....	60
2.6.4 Fog Computing референтна архитектура.....	60

2.7. Fog Computing апликации.....	61
2.8. Иднината на Fog Computing	63
2.9 Примена на пресметувањето во магла кај системите за здравствена нега ..	64
2.9.1 Архитектура на системот и улогата на пресметувањето во магла	68
2.9.2 Својства и карактеристики на паметни порти за е-здравство во магла нивото	72
2.9.2.1 Локална обработка на податоци.....	72
2.9.2.2 Филтрирање на податоци.....	73
2.9.2.3 Компресија на податоци.....	74
2.9.2.4 Фузија на податоци	74
2.9.2.5 Анализа на податоци.....	75
2.9.2.6 Адаптивност	75
2.9.2.7 Локално складирање	76
2.9.2.8 Локално активирање.....	77
2.9.2.9 Безбедност	78
2.9.2.10. Интероперабилност и реконфигурабилност - Interoperability and reconfigurability	79
2.9.2.11. Откривање на уреди и мобилна поддршка	80
2.9.2.12. Енергетска ефикасност на сензорските јазли	82
2.9.2.13 Латентност.....	83
3. Обработка и отстранување на шуми движевски артефакти од PPG сигнали ...	84
3.1. Фотоплетизмографија.....	85
3.2. Движевски артефакти.....	91
3.3. Дефиниција на проблемот	93
3.4. Нашиот алгоритам	98
3.4.1. Алгоритми за декомпозиција на сигналот	100
3.4.2. Сингуларна спектрална анализа (Singular Spectrum Analysis).....	101
3.4.3. Оптимизирање и прилагодување на параметрите	103
3.4.4. Отстранување на движевски артефакти	107
3.4.5. Отстранување на компоненти (RoC).....	110

3.4.6. Идентификација на МА компоненти	111
3.4.7. Резултати од применетиот алгоритам	115
4. Методи на машинско учење	118
4.1 Кластерирање	122
4.2 Методот K-means	122
4.3 Класификација и регресија.....	124
4.3.1 Класична регресија	124
4.3.2 Линеарна регресија (linear regression)	125
4.3.2.1 Регуларна линеарна регресија (Regularized Linear Regression -RLR)	125
4.3.2.2. Риџ линеарна регресија (Ridge Linear Regression - RLR)	126
4.3.2.3 Повеќекратна линеарна регресија (Multiple Linear Regression - MLR)	126
4.3.3 Регресија и класификација на основа на дрва (Regression trees)	127
4.3.3.1 Модел дрва (Model trees)	129
4.3.3.2 Ансамбли на дрва (Ensembles of trees)	129
4.3.3.3 Регресија на пакување (Bagging Regression - BR).....	129
4.3.3.4 Случајни шуми (Random Forest)	129
4.3.4 Машини со носечки вектори (Support vector machine)	132
4.3.5 Баесова класификација	135
4.3.5.1 Наивен Баесов класификатор.....	135
4.3.5.2 Баесова мрежа.....	136
4.3.6 Асоцијативна класификација.....	137
4.3.6.1 RIPPER (JRip).....	138
4.3.7 Симболичка регресија и генетско програмирање	139
4.3.7.1. Поединци и популација	139
4.3.7.2. Фитнес-функција	141
4.3.7.3 Механизам на селекција и генетски оператори.....	142

4.4 Невронски мрежи (Neural Network)	143
4.5 Длабоко учење (Deep learning)	147
4.6 Евалуација.....	151
5. Основни карактеристики што го опишуваат PPG-сигналот.....	153
5.1 Анотацијана PPG сигналите.....	157
5.1.1 Алгоритми за детекција на врвови	158
5.1.1.1 AMPD алгоритам.....	158
5.1.1.2 Алгоритам за анотација на BP	162
5.1.2 Определување на карактеристиките и анотација на сигналот	164
5.2 Влезно податочно множество	166
5.3 Класификација со помош на околината WEKA.....	168
5.4 Резултати од класификација на податоците	171
6. Заклучок.....	175
Користена литература.....	178
Прилог 1	198
Листа на кратенки:	198

Листа на слики

Слика 1.1: Шематски приказ на четирите процеси на податочно рударење (активност, откривање на аномалии и поддршка во поставување на дијагноза) во зависност од различните аспекти на собирање на податоци од сензорските лесно носливи системи (Banaee et al., 2013)	2
Слика 1.2: Типичен ЕКГ сигнал.....	18
Слика 1.3: Типичен PPG сигнал и бран на неговиот втор извод.....	19
Слика 1.4: PTT определено со сигналите ЕКГ и PPG.....	21
Слика 1.5: Приказ на дефиницијата за срцево отчукување преку сигналот ЕКГ ...	22
Слика 2.1: Карактеристики на пресметувањето во магла	30
Слика 2.2: Три слоја на Облак Магла моделот (Mahmood & Ramachandran, 2018)	38
Слика 2.3: Модел на пресметка во магла (Naha et al., 2018)	39
Слика 2.4: Архитектура на поврзаните технологии со Fog Computing (Naha et al., 2018).....	43
Слика 2.5: Fog Computing архитектура на високо ниво (Naha et al., 2018).....	52
Слика 2.6: Компоненти на архитектура на пресметка во магла.....	54
Слика 2.7: Основна архитектура на IoT систем за здравствена нега.....	65
Слика 2.8: Основна IoT архитектура која се заснова на магла	69
Слика 2.9: Составни компоненти на IoT систем за здравствено следење.....	70
Слика 2.10: Паметна архитектура на порта за е-здравство	73
Слика 2.11: Мобилност на јазлите во магла пресметката	82
Слика 3.1: AC и DC компонента на PPG-сигналот	86
Слика 3.2: Типична форма на бранот на PPG-сигналот: а) AC компонента (црвена крива, горната крива); DC (црна крива, испрекинатата) е бавно варијабилна компонента; б) сооднос AC/DC, ЕКГ за истиот сигнал е претставен во сина боја(Kamshilin & Margaryants, 2017)	87
Слика 3.3: Систолна и дијастолна фаза кај PPG-сигналот и дикротичен зарез	88
Слика 3.4: Видови на пулсниоксиметри - режим на пренесување, режим на одбивање	90

Слика 3.5: Периодограм на PPG-сигналот а) Случај кога највисоките вредности на HR и на шумот се наоѓаат на поголемо растојание б) Случај кога врвот на HR е целосно отсутен с) Случај кога врвот на HR не може да се определи	93
Слика 3.6: PPG сигнал со и без движевни артефакти	94
Слика 3.7: График на сопствените вредности.....	106
Слика 3.8: Начините на кои движевните артефакти можат да се појават во PPG сигналите	110
Слика 3.9: Начин на идентификација на компонентите на движевните артефакти	112
Слика 3.10: Шематски приказ на алгоритмот	115
Слика 3.11: Сегменти од резултатите од филтрирањето применети на 3 различни мерења. Првиот дел кај секој од сигналите е измерениот PPG сигнал, вториот е сигнал добиен од првичното филтрирање и третиот е крајно филтрираниот сигнал со отстранетите движевни артефакти.....	117
Слика 4.1: Вештачка интелигенција, машинско учење и длабоко учење.....	118
Слика 4.2: Поделба на машинското учење.....	120
Слика 4.3: Постапкана кластерирање.....	123
Слика 4.4: Дијаграм на тек на MLR (Khalid et al., 2018)	127
Слика 4.5: Дијаграм на тек кај регресивни дрва (Khalid et al., 2018).....	128
Слика 4.6: Разлика во добивката на информации во зависност од избраните карактеристики. Во првиот случај има голема добивка на информации, а во вториот мала	131
Слика 4.7: Претставување на методот за случајни шуми	132
Слика 4.8: Дијаграми на тек кај методот на SVM (Khalid et al., 2018)	133
Слика 4.9: Изразот $((2 + 2) + (2 + 2)) + (3 + 3)$ претставен како дрво на изрази	140
Слика 4.10: Модел на дигитален неврон заснован на моделот McCulloch-Pitts (Bishop, 2006).....	144
Слика 4.11: Пример за едноставна невронска мрежа	145
Слика 4.12: Пример за повеќеслојна feedforward мрежа.....	146
Слика 4.13 500 PPG циклуси каде секој е центриран во прозорец од 2 секунди (Slapničar et al., 2018)	150

Слика 5.1: Приказ на карактеристиките на PPG-бранот во временски домен (Slapničar et al., 2018)	154
Слика 5.2: Морфолошки карактеристики на PPG-сигналот користени во нашиот случај.....	157
Слика 5.3: Постапка за добивање на матрица М (LMS на сигналот x).....	159
Слика 5.4: Тек на AMPD-алгоритамот.....	161
Слика 5.5: Пример од употреба на AMPD-алгоритамот врз симулиран сигнал (Scholkmann et al., 2012)	162
Слика 5.6: Определување на dicrotic notch според (Le Gall et al., 2017)	163
Слика 5.7: Приказ на резултатот со користење на алгоритамот (Laurin, 2017)	164
Слика 5.8: Вредности за крвен притисок кај возрасни.....	171

List of Figures

Figure1.1 Schematic overview of the four main data mining processes (activity, prediction, anomaly detection and diagnose/decision support) in relation to different aspects of wearable sensing in wearable health devices (Banaee et al., 2013)	2
Figure1.2 Typical ECG signal.....	18
Figure1.3 Typical PPG signal and its Second derivative wave.....	19
Figure1.4 PTT determined by the EKG and PPG	21
Figure1.5 An illustration of the definition of Heart Rate (HR).....	22
Figure 2.1 Fog Computing characteristic.....	30
Figure 2.2 Three-layer model of Cloud Fog(Mahmood & Ramachandran, 2018).....	38
Figure 2.3 Fog computing model (Naha et al., 2018)	39
Figure 2.4 Architecture of related Fog Computing technologies (Naha et al., 2018)	43
Figure 2.5 High level architecture of Fog Computing (Naha et al., 2018).....	52
Figure 2.6 Component of Fog Computing architecture.....	54
Figure 2.7 General IoT-based health monitoring system architecture	65
Figure 2.8 Fog-based IoT architecture	69
Figure 2.9 Components of IoT - based health monitoring system	70
Figure 2.10 Smart e-Health gateway architecture	73
Figure 2.11 Node mobility in fog computing	82
Figure 3.1 AC and DC component of PPG signal	86
Figure 3.2 Typical form of PPG signal wave: a) AC component (red curve, the upper curve), DC (black, dotted component is a slowly variable component. b) AC/DC ratio. The ECG for the same signal is present in blue, the bottom curve (Kamshilin & Margaryants, 2017)	87
Figure 3.3 Systolic and diastolic phase in the PPG signal and dicrotic notch.....	88
Figure 3.4 Types of pulse oximeters – transmission mode, reflection mode	90
Figure 3.5 Periodogram of PPG signal. A) Case where the highest values of HR and noise are at a greater distance. b) Case where the top of the HR is completely absent. c) Case where the peak of the HR cannot be determined.....	93
Figure 3.6 PPG with and without moving artifacts	94
Figure 3.7 Graph of eigen values	106

Figure 3.8 The ways when moving artifacts can appear in PPG signals	110
Figure 3.9 Example of moving artifacts identification	112
Figure 3.10 Schematic representation of the algorithm	115
Figure 3.11 Segments from filtering results applied to 3 different measurements. First part of each measurement is measured PPG signal, second part is signal obtained from the initial filtering and third part is filtered signal with the removed moving artefacts ..	117
Figure 4.1 Artificial intelligence, machine learning and deep learning	118
Figure 4.2 Parts of machine learning.....	120
Figure 4.3 Clustering procedure	123
Figure 4.4 MLR flow diagram (Khalid et al., 2018)	127
Figure 4.5 Flow diagram of regression trees (Khalid et al., 2018)	128
Figure 4.6 Difference of gain information depending on selected features. In the first case there is a large gain of information and in the second small	131
Figure 4.7 Presentation of the random forest method	132
Figure 4.8 Flow diagrams of the SVM method (Khalid et al., 2018)	133
Figure 4.9 The expression $((2 + 2) + (2 + 2)) + (3 + 3)$ represented as tree	140
Figure 4.10 Digital neuron model based on the McCulloch-Pitts model (Bishop, 2006)	144
Figure 4.11 Example of a simple neural network	145
Figure 4.12 Example of a multilayer feedforward network.....	146
Figure 4.13 500 PPG cycles, each is centered in a 2 second window (Slapničar et al., 2018)	150
Figure 5.1 PPG wave features in time domain (Slapničar et al., 2018)	154
Figure 5.2 Features of PPG wave used in our case	157
Figure 5.3 Procedure for obtaining matrix M (LMS of the signal x)	160
Figure 5.4 AMPD algorithm flow.....	161
Figure 5.5 Example of using the AMPD algorithm on a simulated signal (Scholkmann et al., 2012)	162
Figure 5.6 Determination of dicrotic notch according to (Le Gall et al., 2017)	163
Figure 5.7 Presentation of the result using the algorithm (Laurin, 2017)	164
Figure 5.8 Blood pressure values for adults	171

Листа на табели

Табела 2.1: Предности на Fog Computing во однос на Cloud Computing	39
Табела 2.2: Разлики меѓу пресметувањето во Магла и пресметувањето на Облак	40
Табела 2.3: Карактеристики на поврзани парадигми со Fog.....	48
Табела 2.4: Разлики меѓу Fog Computing и Edge computing	50
Табела 3.1: Карактеристиката на PPG-сигналот	89
Табела 4.1: Потенцијални градежни блокови на секој поединец (математички израз) во популацијата.....	141
Табела 5.1: Карактеристики на PPG-бранот во временски домен користени кај .	153
Табела 5.2: Карактеристики на PPG-бранот во временски домен користен и во нашиот случај	156
Табела 5.3: Резултати за процент на точно класифицирани податоци од алгоритмите за класификација во WEKA	172
Табела 5.4: Резултати за процент на точно класифицирани податоци од алгоритмите за класификација во WEKA, со помал број влезови	173

List of Tables

Table 2.1 Advantages of Fog Computing over Cloud Computing	39
Table 2.2 Differences between Fog and Cloud computing.....	40
Table 2.3 Characteristics of related Fog paradigms.....	48
Table 2.4 Differences between Fog Computing and Edge computing	50
Table 3.1Characteristic of PPG signal.....	89
Table 4.1 Potential building blocks of each individual (mathematical expression) in the population.....	141
Table 5.1 PPG wave features in time domain used in (Slapničar et al., 2018).....	153
Table 5.2. PPG wave features in time domain, used in our case	156
Table 5.3 Results for percentage of accurately classified data from WEKA classification algorithms.....	172
Table 5.4 Results for percentage of accurately classified data from WEKA classification algorithms, with a smaller number of inputs	173

Резиме

Биофизиолошките параметри кои континуирано можат да се следат со сензори, можат да дадат јасна слика за општата здравствена состојба на човекот во кој било момент. Таквите биофизиолошки параметри се: електрокардиограм (ЕКГ), брзината на срцеви отчукувања (heart rate-HR), крвен притисок (blood pressure-BP), брзината на дишење (respiration rate-RR), крвна сатурација со кислород (Blood Oxygen Saturation-SpO2), шеќер во крвта, потење на кожата, телесна температура и други. Од биофизиолошките параметри кои можат континуирано да се следат со помош на сензори, најпредизвикувачки е крвниот притисок, чие континуирано следење сè уште е во фаза на истражување. За неговото следење се потребни дополнителни сигнали преку кои може да се процени неговата вредност.

Висок или низок крвен притисок или неговите брзи флукуации можат да бидат поврзани со разни болести или состојби. Во денешно време, високиот крвен притисок е еден од најчестите здравствени проблеми во светот. Високиот крвен притисок може да претходи на сериозни срцеви заболувања, мозочен удар и на откажување на бубрезите. Точното мерење на крвниот притисок е од витално значење за дијагностицирање, превенција и третман на овие болести. Крвниот притисок обично се мери во болниците, како дел од вообичаена медицинска рутина.

Сепак, постои зголемена побарувачка за методологии, системи, како и точни и ненаметливи уреди што ќе овозможат континуирано мерење на крвниот притисок и мониторинг на широк спектар на пациенти, овозможувајќи им да ги извршуваат своите дневни активности без никакви нарушувања. Технолошките достигнувања во последната деценија создадоа можности за употреба на разни методологии и уреди како дел од амбиентално живеење за подобрување на квалитетот на животот на луѓето во нивната природна средина.

Во овој докторски труд е развиен систем со иновативна архитектура за континуирано следење на биофизиолошките параметри, посебно на крвниот притисок за навремено алармирање при некои брзи промени и преку тоа да се

обезбеди навремена и соодветна реакција, избегнување и спречување на некои сериозни ситуации кои можат да бидат со катастрофални исходи.

Овој докторски труд предлага архитектура во која податоците се пред-процесираат во мрежните уреди на рабовите од мрежата и на тој начин ја подобруваат ефикасноста и перформансите на самиот систем, истовремено намалувајќи го мрежниот сообраќај и латентноста.

Системот е составен од сензори за собирање на податоци кои лесно можат да се носат и да не пречат во секојдневните рутини, истовремено е направена имплементација на алгоритми базирани на вештачка интелигенција, за обработка на сигнали, за да се направи деривација на потребните биофизиолошки параметри во реално време. На тој начин може да се овозможи и брза интервенција кога е потребно.

Клучни зборови: крвен притисок, PPG сигнал, магла пресметка, машинско учење.

Abstract

Biophysiological parameters that can be continuously monitored with sensors can give a clear picture of a person's general health, at any time. Such biophysiological parameters are: electrocardiogram (ЕКГ), heart rate (HR), blood pressure (BP), respiration rate (RR), blood oxygen saturation (SpO₂), blood glucose, skin perspiration, body temperature and others. Of the biophysiological parameters that can be continuously monitored with the help of sensors, the most challenging is blood pressure, whose continuous monitoring is still in the research phase. To be precisely monitored, additional signals are needed.

High or low blood pressure or its rapid fluctuations can be associated with various diseases or conditions. High blood pressure is one of the most common health problems in the world nowadays. High blood pressure can be the reason for serious heart disease, stroke and kidney failure. Accurate blood pressure measurement is vital for the diagnosis, prevention and treatment of these diseases. Blood pressure is usually measured in hospitals as part of a medical routine.

However, there is an increased demand for methodologies, systems, as well as devices that will allow continuous measurement of blood pressure and monitoring of a wide range of patients, allowing them to perform their daily activities without any inconveniences. Technological advances in the last decade have created opportunities for use of various methodologies and devices as part of an ambient assisted living, in order to improve the quality of life of people in their natural environment.

In this doctoral dissertation a system for continuous monitoring of biophysiological parameters, especially BP is developed. The aim of the system is alerting in case of sudden changes, in order to provide appropriate response, and avoid or prevent some serious situations, which can have catastrophic outcomes.

This doctoral thesis proposes an innovative Fog architecture in which the data is pre-processed in the network edges, and thus the efficiency and performance of the system itself is improved, while network traffic and latency is reduced.

The system consists of sensors for data collection that can be worn and not interfere with daily routines. At the same time algorithms based on artificial intelligence,

for signal processing are implemented, in order to derive the necessary biophysiological parameters in real time. Therefore, rapid intervention is possible when needed.

Keywords: blood pressure, PPG signal, Fog computing machine learning

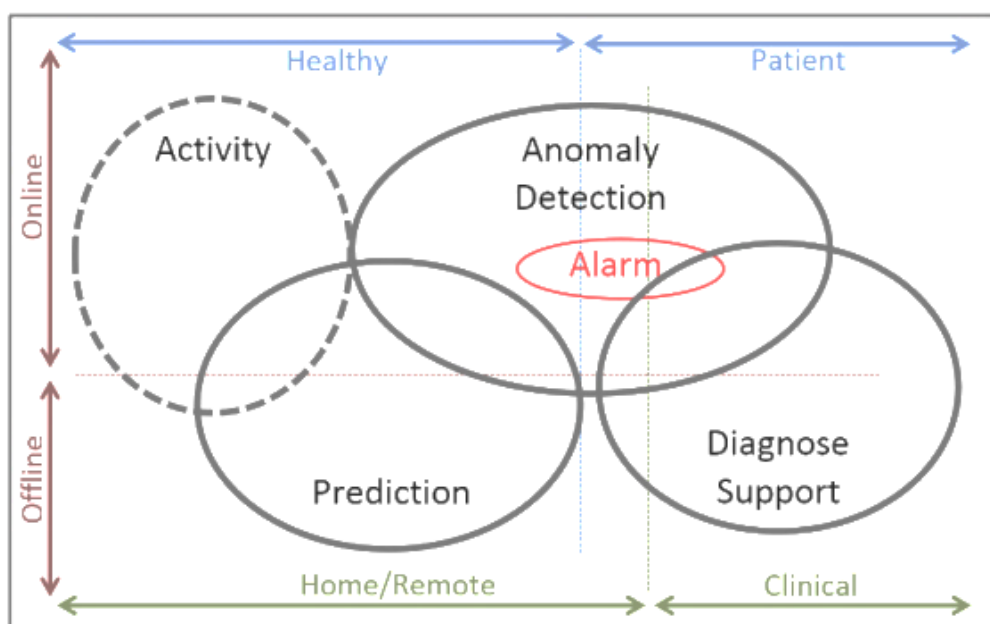
1. Вовед

Континуираното следење на биофизиолошките параметри кај човекот, во амбиентални услови, е од големо значење и претставува фокус на денешната медицина. Ваквиот начин на следење може да придонесе за навремено откривање на голем број здравствени проблеми и состојби, може да помогне во следењето и третманот на пациентите. Денес овој процес е реалност благодарение на технолошкиот напредок во повеќе сектори, почнувајќи од развојот на иновативни материјали, преку микроелектронски компоненти, па сè до развојот на информатичко-комуникациските технологии. За оваа намена, денес се користат системи базирани на сензори кои се ненаметливи, лесно се носат и не претставуваат пречка во нормалното извршување на секојдневните човекови активности.

1.1 Биофизиолошки параметри

Денес постојат голем број сензори што можат да се закачат на едноставен и неинвазивен начин на човековото тело, можат лесно да се носат и да овозможат собирање на различни витални биофизиолошки параметри односно сигнали. Ваквите сензори овозможуваат следење на здравствената состојба на човекот и надвор од клиничката околина. На овој начин се овозможува следење на виталните сигнали за време на различни дневни активности, со што се овозможува подобар увид во здравјето на човекот, како и поставувањето на медицинска дијагноза или подобрување на процесот на закрепнување после интервенции или повреди. Ваквиот начин на континуирано следење на биофизиолошките параметри е корисен и при извршувањето на спортски активности, па дури и во случаи на обука за воени активности кога се следи одговорот на телото кај војникот во различни опасни ситуации, преку што полесно може да се превенира доколку се забележи некоја отстапка во здравствената состојба (Achkoski et al., 2017; Kocev et al., 2018; Stevanoski et al., 2015, 2016).

Човековото тело има повеќе различни биофизиолошки параметри што се мерливи, од електрични до биохемиски. Човековите биофизиолошки параметри можат да се извлечат и да се користат за подобро разбирање на телесното здравје и неговата реакција на надворешните фактори. Затоа, од големо значење е да се разберат главните биофизиолошки параметри што придонесуваат за подобра здравствена анализа на човечкото тело. Сензорските системи за собирање и следење на биофизиолошките параметри можат да се поделат според три аспекти (слика 1.1). Сценарио на употреба (домашна/далечински или клиничка околина), начинот на следење (офлајн или онлајн) и типот на корисници (здави лица или пациенти) (Banaee et al., 2013).



Слика 1.1: Шематски приказ на четирите процеси на податочно рударење (активност, откривање на аномалии и поддршка во поставување на дијагноза) во зависност од различните аспекти на собирање на податоци од сензорските лесно носливи системи (Banaee et al., 2013)

Figure 1.1 Schematic overview of the four main data mining processes (activity, prediction, anomaly detection and diagnose/decision support) in relation to different aspects of wearable sensing in wearable health devices (Banaee et al., 2013) .

Според ваквата класификација, видливи се две различни области, област за набљудување на активности (1) и медицинска област (2), поделена на поткатегории (слика 1.1).

(1) Активност - во која се вклучени апликации за фитнес/здравствена благосостојба други апликации што не се медицински за „лично следење“ и рехабилитација.

(2.1) Предвидување - се состои од идентификација на настани коишто уште не се случиле, обезбедувајќи медицински информации за превенција на идни хронични проблеми, а некогаш може да помогне и во поставувањето на дијагнози.

(2.2) Детекција на аномалии - дел кој е одговорен за идентификување на невообичаени настани/ситуации, кој не е составен дел на нормалното поведение базирано на методи на класификација преку кои се разликуваат нормалните текови од останатите. Алармирањето е составен дел на детекцијата на аномалиите. При што секогаш кога ќе се забележи аномалија се врши алармирање.

(2.3) Поддршка во процесот на поставување на дијагноза – е еден од главните делови во клиничкото следење, кое резултира во донесување на одлука според знаењето откриено од виталните сигнали, здравствените записи и откриените аномалии.

Биофизиолошките параметри кои можат континуирано да бидат следени како витални сигнали се: електрокардиограм (ЕКГ), брзината на срцеви отчукувања (heartrate-HR), крвен притисок (bloodpressure-BP), брзината на дишење (respirationrate-RR), кислородна сатурација на крвта (BloodOxygen Saturation-SpO2), шеќер во крвта, потење на кожата, телесна температура итн. (Dias & Cunha, 2018).

1.1.1 Електрокардиограм

Електрокардиограм (ЕКГ-Electrocardiogram) е еден од најчесто користените биофизиолошки параметри, кој се користи како дијагностичка алатка во здравствена околина. Тој обезбедува информации за срцевиот електричен пулс. Бранот на сигналот на ЕКГ (ЕКГ) се карактеризира со 5 врвови (наречени P, Q, R,

S, T, U). Секој од нив претставува промена на електричниот потенцијал на срцето што кореспондира на соодветна мускулна активност. Најизразениот врв на ЕКГ е врвот R вклучен во комплексот QRS кој претставува деполяризација на коморите каде има поголем потенцијал за диференцијација. Поради оваа причина, последователните R-врвови (интервал R-R) се користат за мерење на срцевите циклуси. ЕКГ - брановата форма се користи за да се анализира срцевиот ритам и да се предвидат и третираат акутни миокардни повреди и коронарни настани.

Анализата на моделите на брановата форма на ЕКГ игра голема улога во дијагностицирање на кардиоваскуларните заболувања, како што се атријална фибрилација, ангина, атеросклероза, срцева аритмија, конгестивна срцева слабост, коронарна артериска болест, срцев удар или тахикардија (P. J. Xu et al., 2008). Една од предностите на континуираното следење на ЕКГ, е подобрувањето на раното откривање на срцевите аномалии заради долги периоди на постојаниот мониторинг, споредено со единичен 24-часовен холтер. Холтерите се уреди што обично се користат за 24-часовно следење на ЕКГ, а нивниот главен недостаток е фактот дека холтерот ја прекинува секојдневната рутина на животот и тоа го прави некорисен за ненаметлив континуиран мониторинг. За да се надмине овој проблем, развиени се системи преку кои сензорите се вградени во ткаенини со употреба на различни материјали (Yilmaz et al., 2010; Xu et al., 2008). Овој вид на сензори не предизвикува иритација на кожата, но, бидејќи тие не се залепени за кожата, артефактите се појавуваат како резултат на движењето на телото, што ги прави, клинички, потешки за обработка. За да се намалат артефактите произведени од движењето на телото и иритација на кожата, (Luo et al., 2014) развиле нова технологија на суви, флексибилни сензори кои можат да бидат прикачени за човечката кожа.

1.1.2 Срцево отчукување (heartrate-HR)

HR како биофизиолошки параметар се користи и во здравството и во фитнес, односно во спортските активности. Следењето на овој сигнал обезбедува информации за физиолошкиот статус со индицирање на промени во срцевиот циклус. Овој витален параметар може лесно да се извлече од ЕКГ сигналите (преку мерење на растојанието меѓу R-врвовите) или од фотоплетизмографот

(PPG сигнал) (Chan et al., 2012). Иако овие два биофизиолошки параметра имаат различни морфолошки информации во нивните бранови форми и се од различно физиолошки потекло, тие содржат слични информации за отчукувањата на срцето. Постојат и други начини за мерење на брзината на отчукувања на срцето, на пример, користење на инерцијални сензори (Aarts et al., 2017) или скали (Giovangrandi et al., 2012), наречени балистокардиограм (BCG), но тие се методи кои немаат изводливо мерење во споредба со HR добиено од ЕКГ и PPG. За време на спортски и други поинтензивни активности, од голем интерес е да се открие како ќе варира отчукувањето на срцето. Анализата на варијабилноста на срцевиот ритам, привлекува внимание бидејќи претставува едноставен показател за здравствената состојба на кардиоваскуларниот систем. Исто така, тоа е и индикатор за човековиот психофизиолошки статус, бидејќи различно реагира/варира при стрес и напор (Ahrens, 2008; Teng et al., 2008). Сличен на овој биофизиолошки параметар е пулсниот сигнал кој е поврзан со брзината на отчукување на срцето и може да го замени него. Пулсниот сигнал е дефиниран како опиплива ритмичка експанзија на артерија, произведена од зголемување на волуменот на крвта туркана во садот, предизвикана од контракција и релаксација на срцето. Ваквото мерење овозможува повеќе информации како што е јачината, амплитудата и регуларноста на пулсот. Пулсниот сигнал не треба да се поистоветува со срцевиот ритам и може да се добие со употреба на принципи на пулс оксиметрија, метод што, исто така, се користи за мерење на заситеноста на кислородот во крвта (Xu et al., 2008).

1.1.3 Крвен притисок

Крвниот притисок (Bloodpressure - BP) се смета за најважен кардиопулмонален параметар, што укажува на притисок извршен од крв врз артерискиот ѕид. BP обезбедува индиректни информации за протокот на крвта кога срцето врши контракција (систола) и се релаксира (дијастола) и може да укаже на испораката на клеточен кислород. Овој биофизиолошки параметар е под влијание на неколку човечки физиолошки карактеристики: срцев исход, периферна васкуларна отпорност, волумен на крвта и вискозност и еластичност на ѕидот на садовите. Амбулантското следење на BP овозможува да се добијат испитувања за крвниот

притисок неколку пати на ден, што е идеално за следење на високиот крвен притисок (хипертензија), една од најголемите закани за повеќе болести, како и за подобрување на предвидувањето на кардиоваскуларните заболувања (Lameski et al., 2017; Elliott & Coventry, 2012; Turner et al., 2015;). Крвниот притисок, традиционално се мери со употреба на манжетни под притисок со стетоскоп на раката на пациентот. Овој метод е адаптиран за да се изврши автономно мерење на БП, вклучително и со користење на целосно автоматизирана манжетна за наддување што го мери крвниот притисок, со поврзување на надворешниот притисок со големината на пулсирањата на артерискиот волумен (Yilmaz et al., 2010). Постојаното носење на манжетна може да резултира во несакани ефекти, како што се нарушување на спиењето, иритации на кожата и зголемување на нивото на стрес. Затоа, развојот на нови технологии и методологии за континуирано следење на крвниот притисок, денес претставува голем научен предизвик и мотив за голем број на научни истражувања (Dias & Cunha, 2018).

1.1.4 Фреквенција на дишење (RespirationRate -RR)

RR е основен биофизиолошки параметар при следење на здравствената состојба на пациентот. Кај критични заболувања, ова е еден од најчувствителните показатели, како што е во случај на вознемиреност, а може да биде и најава за други несакани настани, како што е срцев удар. Амбулантското следење на фреквенцијата на дишење е важно при откривање на симптоми на респираторни заболувања, како што се синдром на ноќна апнеа, хронично опструктивно белодробно заболување, астма и слично. Овој постојан мониторинг е особено важен кај деца со белодробни заболувања (Chan et al., 2012; Elliott & Coventry, 2012).

Овој витален параметар обично се пресметува од добиената респираторна бранова форма што ја рефлектира варијацијата на волуменот на градниот кош за време на вдишување и издишување. Анализата на овие податоци во спортот, може да помогне во постигнување на подобри респираторни перформанси (Chan et al., 2012; Elliott & Coventry, 2012; Teng et al., 2008). Во денешно време, за да се добие респираторната функција, постојат три основни методи: еластомерна плетизмографија (elastomeric plethysmography - EP), импедансна плетизмографија

(impedance plethysmography-IP) и респираторна индуктивна плетизмографија (respiratory inductive plethysmography - RIP).

Техниката EP ја претвора тековната варијација на пиезо-електричните сензори во напон со помош на еластичен ремен. (Guo et al., 2013) Гуо и останатите во 2013 г. развиле прототип за облека што може да ги мери промените во волуменот на градниот кош и стомакот со голема точност, користејќи сензор за пиезорезистентна ткаенина.

IP користи промени на импеданса на површината на телото како резултат на проширување и контракцијата за време на дишењето. Оваа технологија била користена во развојот на униформа што треба да се користи кај војниците (Seoane et al., 2014). Принципот на технологија RIP се заснова на јамка со жица на струја што генерира магнетно поле нормално на ориентацијата на јамката. Варијациите на волуменот на градниот кош ја менуваат областа затворена со јамка, создавајќи спротивна пропорционална струја (Mazeika & Swanson, 2007).

Покрај овие три основни методи, други технологии кои се користат за да се добие респираторната бранова форма се: акцелерометри (Jin et al., 2009); извлекување од сигналот за ЕКГ (Sharma et al., 2015) добивање од пулсова оксиметрија (Addison et al., 2014); полимерни сензори (Chiu et al., 2013); оптички влакна (Krehel et al., 2014); итн.

1.1.5 Кислородна заситеност (Blood Oxygen Saturation - SpO2)

Кислородната заситеност на крвта (SpO2) е исклучително важен витален параметар и е лесен за мерење со употреба на технологија за фотоплетизмографија (PPG) и принципи на пулснаоксиметрија. Методот PPG овозможува да се добие брановиот облик на варијација на крвните садови со користење на две бранови должини (660 nm и 905 nm). На тој начин можно е да се процени кислородната заситеност на крвта. Ова се должи на промената на спектарот на апсорпција на хемоглобин кога граничи со кислород. Користејќи оксиметрија, може да се процени количеството кислород што го носат крвните клетки (нормално: 95-100%). Оваа мерка може да доведе до откривање на промена на состојбата на пациентот што во спротивно може да се пропушти, како што е помал процент на кислород (<95%), што укажува на недоволно снабдување

со кислород во човечкото тело. Еден од проблемите во мерењето на заситеноста на кислородот во крвта е кога пациентот е анемичен (Dias et al., 2017; Elliott & Coventry, 2012; Tamura et al., 2014).

Покрај медицинската употреба, следењето на кислородната сатурација има посебен интерес за проценка на аеробната ефикасност на лице кое врши рутинска вежба. Информациите за заситеноста со кислород во екстремитетите и мозокот се исто така важни во воените и вселенските апликации, во кои гравитациските промени можат да влијаат на испорака на кислород до овие делови од телото што доведува до губење на свеста. Исто така, откриено е и постоење на позитивна корелација помеѓу перформансите на поединецот во извршување на некоја задача и присуството на кислород (Teng et al., 2008).

Постојат неколку неинвазивни технологии за мерење на заситеноста со кислород, кои можат да се прилагодат на уреди што се носат. Методот со користење на PPG се истакнува како покорисен (Tamura et al., 2014). PPG-сигналот за добивање на нивото на заситеност со кислород во крвта, најчесто се добива од прстот.

Ушниот лобус, исто така, може да се користи, а неодамнешните истражувања покажуваат постоење на многу мал чип (3×6 мм) способен за мерење на заситеноста со кислород на крвта. PPG сензорите во челото се користат за мерење оксигенација на мозокот (Mendelson et al., 2013). А исто така, за оваа намена направени се и повеќе обиди за поставување на PPG сензори на облека.

1.1.6 Шеќер во крвта

Шеќерот во крвта е биофизиолошки параметар чие следење, особено е значајно кај дијабетичарите. Болеста на дијабетес предизвикува неколку физиолошки нарушувања, како на пример, нарушување на церебралното и на васкуларното ткиво, ретинопатија и нефропатија. За да се спречи тоа, лицата со дијабетес ја контролираат концентрацијата на шеќерот во крвта, со постојано мерење и инјектирање на инсулин, кога е потребно, а за да се одржуваат нормалните вредности. Најкористен метод за проценка на концентрацијата на шеќер во крвта е земање на примерок од крв со боцкање во прстот.

Некои примери за континуирано мерење на шеќер во крвта се Medtronic Continuous Glucose Monitoring (CGM) (Medtronic Diabetes, 2015), Dexcom G4 Platinum (Garcia et al., 2013) уреди кои се способни за мерење на нивото на шеќер со помош на леплива лепенка со игла.

Направени се повеќе напори за создавање на уреди преку кои со неинвазивни методи можат да се мерат нивоата на шеќер во крвта (Appelboom et al., 2014; Takahashi et al., 2013). Еден од првите комерцијални уреди што се носат за континуирано следење на гликозата во крвта е Gluco Watch®, кој овозможува да се мерат нивоата на шеќер во крвта на секои 20 минути во текот на 12 часа преку кожата со јотофереза (Tierney et al., 2001). Во 2007 година е повлечен поради ефектот на горење на кожата. Други неинвазивни техники, што се развиени подоцна се: биоимпедантна спектроскопија (So et al., 2012), електромагнетни сензори кои се многу осетливи на температура. Исто така, други обиди се мерење на шеќер во крвта преку око и користење на ултразвук (Sobel et al., 2014).

1.1.7 Потење на кожата

Потењето на кожата не е клинички параметар, туку биофизиолошки знак што се користи за анализа на човековата реакција на неколку ситуации. Одредени ситуации можат да предизвикаат невролошки реакции од автономниот нервниот систем (ANS) со што се стимулира зголемување на потењето на кожата. Влагата ја менува електричната спроводливост на кожата, овозможувајќи мерење на количината на пот, произведена од потните жлезди, именувана како галвански одговор на кожата (galvanic skin response - GSR). Бидејќи ANS е одговорен да контролира други физиолошки параметри како отчукувањата на срцето, дишењето и крвниот притисок, GSR се користи заедно со собирањето на некои од овие сигнали. На пример, потењето на кожата и варијабилноста на срцето можат да се користат за класификација на менталните состојби, на тој начин помага во разграничувањето како и во откривањето на менталниот стрес (Jeehoon Kim et al., 2014; Nikolic-Popovic & Goubran, 2011).

Исто така и во спортот, постојаното следење на потењето на кожата се смета за значаен биофизиолошки знак, особено во превенција на дехидратацијата (Chan et al., 2012; Nikolic-Popovic & Goubran, 2011).

Со помош на потењето на кожата можат да се идентифицираат некои патолошки нарушувања преку јонските нивоа во составот на флуидот. Нивоата на натриум, амониум и калциум се индикатори за нерамнотежа на електролити, но исто така и на цистична фиброза, остеопороза, губење на минерали во коскитеи физички стрес. На пример, физичкиот стрес може да се користи при психофизиолошка проценка на војската која е подложна на интензивна обука (Bandodkar & Wang, 2014).

Постојат два главни вида на сензори при следењето на потењето на кожата: епидермално базирани сензори и флексибилни пластични сензори. Првите директно се прикачени на епидермисот на кожата, а вторите се вградени во ткаенината или прикачени на неа (Dias & Cunha, 2018).

1.1.8 Капнографија (Capnography)

Капнографијата е неинвазивен и економичен метод за проценка на вентилацијата кај луѓето, што укажува на нивото на јаглерод моноксид присутен во циклусот на дишењето, што е многу корисно за да се избегнат различни клинички проблеми.

Капнографијата континуирано го мери вдишаниот и издишаниот парцијален притисок на јаглерод диоксид (PCO_2) во респираторните гасови. Ова мерење е направено преку фаќање на воздухот веднаш под носот, каде што се поставува уредот за капнографија за да изврши квантификација на гас на CO_2 , добивајќи карактеристична бранова форма од која, исто така, може да се добие брзината на дишење (Kodali, 2013; Wac & Tsiourti, 2014).

Повеќе од 25 години, капнографијата има широка употреба во клиничката пракса во операционите простории при анестезија, дозволувајќи им на анестезиолозите да го оценат нивото на свеста на пациентот. Капнографијата може да се користи надвор од клиничкото опкружување за да се следи синдромот на апнеа при спиење.

Се очекува, капнографијата да стане пошироко користена и надвор од клинички средини. За да се овозможи ова, потребно е да се направат сигурни,

економични и преносни единици за капнографија со брзи процедури за калибрација (Dias & Cunha, 2018).

1.1.9 Телесна температура

Температурата на телото е резултат на балансот помеѓу производството на топлина и загубата на топлина во организмот. Нејзиното следење е од суштинско значење за да се избегне дефункционализација на многу елементи заради високи температури. Телесната температура се дели на два дела: јадрена (основна) температура и температура на кожата. Температура на кожата варира во поширок опсег на температури од основната температура. Температурата на кожата е под влијание на циркулацијата на крвта и исто така е поврзана со HR и со метаболички процеси. Надворешни фактори, како што е циркулацијата на воздухот, температурата на околината и влажноста, исто така играат важна улога во овој механизам за регулација на телесната температура (Gaura et al., 2013).

Развиени се различни системи што се носат за мерење на двата вида на температурата, како на пример температурни сензори залепени за кожата кои континуирано се носат. Сепак, сè уште остануваат предизвик, неинвазивните начини за мерење на јадрената температура преку отчукувањето на срцето и мерењето на температурата на кожата. Тоа главно се должи на надворешните фактори кои можат да влијаат на физиолошките знаци, со што е тешко да се има директна корелација помеѓу овие променливи кои зависат само од човековиот механизам за физиолошка терморегулација (Buller et al., 2010).

1.2 Континуирано следење на крвниот притисок

Крвниот притисок, кој може да се дефинира како притисок направен со циркулирање на крвта врз сидовите на крвните садови, претставува значаен биомаркер на кардиоваскуларното здравје и индикатор за општо здравје (Klabunde, 2012). Крвниот притисок се смета за еден од најважните витални знаци на кардиоваскуларниот систем (Center for Disease Control and Prevention, 2011). Тој може да варира поради некоја физичка активност, вознемиреност, лекови и различни емоции. Флуктуациите на крвниот притисок може да бидат знак на некое

здравствено нарушување. Поопасни за здравјето се флукуациите кога крвниот притисок се зголемува. Хипертензијата, состојба на висок или зголемен крвен притисок, претставува глобален здравствен проблем. Високиот крвен притисок, во денешно време претставува најчеста причина за смрт и инвалидитет во светот, а исто така е главен фактор на ризик за многу сериозни болести, како што се заболувања на бубрезите и кардиоваскуларни проблеми. Високиот крвен притисок, исто така, може да биде причина за коронарна артериска болест и различни срцеви заболувања. Хипертензијата може да предизвика хронично заболување на бубрезите, а исто така може да биде фактор на ризик за проблеми со артериите, мозочен удар или периферна артериска болест (Mitchell, 2014; Rosendorff et al., 2015). Според извештајот на Светската здравствена организација (СЗО), покачениот крвен притисок во моментов директно или индиректно е причина за смрт на девет милиони луѓе годишно (World Health Organization, 2013). Хипертензијата ретко предизвикува симптоми во раните фази и затоа, повеќето пациенти со хипертензија обично не се свесни за постојната болест. Оние кои се дијагностицирани обично немаат никаков третман во рана фаза и затоа не се во можност успешно да ја контролираат својата болест долгорочно. Иако тие не се ни свесни за покачениот крвен притисок, тие тивко ги убиваат нивните внатрешни органи на телото (на пример бубрезите, мозокот и очите), поради што хипертензијата е позната и како тивок убиец. Дневното мерење и следење на крвниот притисок е клучно за рано откривање на здравствени проблеми, особено на проблеми со срцето и мозочен удар (Halдар, 2013).

Крвниот притисок обично е претставен со три вредности: систолен (systolic blood pressure - SBP), дијастолен (diastolic blood pressure - DBP) и среден притисок (mean blood pressure - MBP). Систолен притисок е притисок направен на сидот на крвните садови со крв на крајот од систолната контракција на коморите. Систолниот е максимален притисок. Се јавува кога крвта се пумпа од левата комора во аортата. Дијастолниот притисок е притисок направен на сидовите на крвните садови на крајот на релаксација или дијастола. Дијастолниот е минимален притисок. Се јавува кога крвта тече од преткоморите во коморите. Средниот притисок е притисок што ја одредува просечната брзина на протокот на

крв низ системските садови. Средниот притисок обично се пресметува со следнава формула: $MBP = 1/3 \times (SBP + 2 \times DBP)$.

1.2.1 Можни методи за мерење на крвниот притисок

Постојат два можни пристапа за мерење на крвниот притисок: инвазивни и неинвазивни. Инвазивниот систем за мерење и следење на крвниот притисок користи катетер, односно тенка флексибилна цевка што се вметнува во артеријата. Поточно, читањето на крвниот притисок може да се добие со употреба на инвазивни мерења. Оваа техника најчесто се користи за следење на крвниот притисок кај високоризични хируршки пациенти и пациенти во единицата за интензивна нега. Инвазивното следење на крвниот притисок обично се прави преку канилација на периферна артерија со катетер. Во податоците, кои се добиваат со инвазивно мерење, освен само податоци за крвниот притисок се наоѓа и голема количина на други дополнителни информации. Овој вид мерење, за да може да се организира и изведе има потреба од постојана контрола на експерти, лекари и медицински лица. Според тоа, овој начин на мерење не е соодветен за употреба надвор од болнички рамки (Chung et al., 2013)

Најчестиот неинвазивен метод за мерење на крвниот притисок, користен уште во минатиот век е со поставување на апаратот за мерење на горниот дел од раката. Овој пристап може да предизвика пациентите да се чувствуваат непријатно и да чувствуваат напнатост во раката, а сето тоа може да придонесе до дополнителен стрес, што може да доведе до грешки во мерењето. Неинвазивно мерење на крвниот притисок, обично се прави со употреба на живин сфигмоманометар и осцилометар, но тие главно се засноваат на воздушна манжетна, што само може да го мери крвниот притисок на периоди, така што ваквиот начин на мерење не може да биде соодветен за долгорочно следење на крвен притисок.

Стандардниот начин за мерење на артериски ВР со жиг сфигмоманометар се користи повеќе од 100 години. Ова е најчестиот метод на мерење на ВР што се користи и во денешно време. Сфигмоманометар со жива може да работи на начин заснован на дување на манжетна околу раката на пациентот, за да се спречи протокот на крв во артеријата. Додека манжетната полека се отклонува,

звучите од Короткоф се откриваат со стетоскоп поставен над брахијалната артерија на пациентот. Овој метод бара обучен персонал за откривање на овие звуци. Притисоците на наметките на кои е откриен првиот звук Короткоф го претставуваат систолниот, а притисоците на манжетната на кои е откриен петтиот звук Короткоф е дијастолниот притисок. Оваа техника е позната како аускултативна техника. ВР може лесно да се смени и може да варира секоја минута. Затоа, неопходно е резултатите од мерењето на ВР да се стават во графикон, па со текот на времето да се добие точна проценка. Од таа причина, во клиниката за управување со хипертензија, неопходно е континуирано 24-часовно следење на ВР. Ова е поинтензивно користено.

Осцилометријата е уште еден начин за неинвазивно мерење на крвниот притисок и се користи за секојдневно набљудување на клиниките во болниците и центрите или домовите за нега. Овој вид на мерење користи наметка за надувување на раката или на зглобот на пациентот за да го спречи протокот на крвта. Со овој метод не се откриени Korotkoff звуци, но наместо тоа, кога наметката полека се отклонува, се користи трансдуктор за да се сними осцилацијата на притисокот. Притисокот, кога се појавува максимална осцилација е средниот притисок (МВР), а систолниот и дијастолниот подоцна се пресметуваат од добиениот МВР.

Тонометријата е уште еден неинвазивен континуиран метод на мерење на крвниот притисок, исто така познат како „васкуларно истоварување“, во кое низа сензори се притискаат врз кожата над артеријата. Овој метод што се користи во клинички услови е способен да обезбеди точна снимка на артериски бранови, но има две главни ограничувања: поставеноста на тонометар над артеријата е многу критична и е многу чувствителна на движење. Бидејќи тонометријата се применува на периферната циркулација, обично се користи алгоритам за проценка на точниот крвен притисок врз основа на снимениот. Тонометрискиот сензор мора да биде заштитен од движење и други механички надворешни промени за да се добие стабилен сигнал (Chung et al., 2013).

Волуменозни стегач (volume-clamp) е уште еден метод за неинвазивно континуирано мерење на крвниот притисок, што го мери артерискиот притисок на

прстите со помош на манжетна за прсти и меур за надувување во комбинација со инфрацрвен плетизмограф. Плетизмограф е уред за мерење на волуменот кој содржи инфрацрвен извор на светлина и детектор. Крвта апсорбира инфрацрвена светлина, но пулсирањето на артерискиот сид за време на чукање на срцето предизвикува пулсирање во сигналот на детекторот на светлина. Finapres (од медицински системи Finapres) е таков вид на уреди и се користи и во денешно време. Иако, овој метод обично генерира добри резултати од мерењето и користи помала манжетна, сепак поради амбулантните карактеристики на уредот за обработка сè уште го прават несоодветен за сигурно и удобно носење преку целиот ден, а со тоа и не соодветен за обезбедување континуирани податоци. Покрај тоа, ваквиот начин е чувствителен на движење и не може да се користи за мерење на крвниот притисок при нормални дневни активности (Ma, 2014).

Технолошките достигнувања во последната деценија создадоа можности за огромен број на апликации, со амбиентално помогнато живеење (Ambient Assisted living - AAL) (Chandrasekaran et al., 2013; Koceska et al., 2017; E. T. McAdams et al., 2010; Silva et al., 2014) во различни полиња, што можат да имаат корист од континуираното следење на крвниот притисок. Меѓутоа, за да се провери крвниот притисок на една личност за време на неочекувани ситуации, во секое време, на кое било место и да се добијат континуирани мерења, потребен е уред или апарат кој ќе биде преносен, удобен и лесен за носење. апаратот не треба да му штети на корисникот во секојдневните рутини. Неодамнешните достигнувања во вградените системи, мобилните компјутери и безжичните мрежи овозможуваат постојан мониторинг на крвниот притисок, дури и надвор од болниците. Новоразвиените методи, електронските системи, производите и услугите кои транспарентно им помагаат на луѓето во нивното секојдневно живеење, можат еноормно да го охрабрат амбиенталното живеење како начин за подобрување на квалитетот на животот во сите фази на животот кај луѓето во нивната природна средина. Овој нов пристап може да го подобри управувањето со хроничните заболувања. Ова може да се направи со поттикнување на промени во животниот стил и ефективно рано откривање и лекување на многу здравствени проблеми пред да им треба скапа и итна интервенција. Затоа, постои итна потреба од нови

системи за набљудување кои можат да вклучуваат нови технологии на сензори, алгоритми и амбиентална интелигенција кои се способни да се справат со следење на пациентите. Сето тоа мора да биде направено, дискретно во домовите на пациентите без значително да се наруши нивната удобност, додека тие ги извршуваат нивните секојдневни активности.

1.2.2 Методологии за континуирано мерење на крвниот притисок

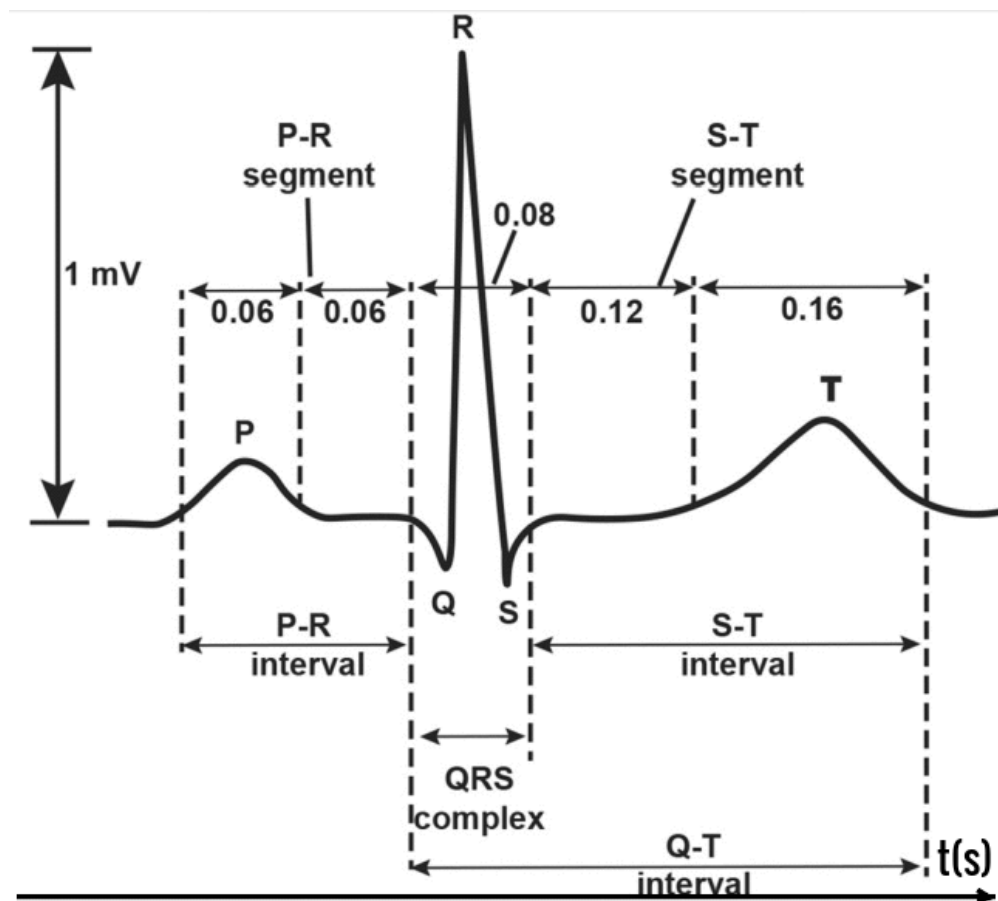
Со напредокот во обработката на дигиталните сигнали и зголемената компјутерска моќ, а исто така и со зголемената потреба за производи за следење на здравјето, значително внимание добиваат компјутерската (автоматска) проценка на крвниот притисок и новите сензори што лесно се носат и можат да се искористат за континуирани неинвазивни решенија за мерење на крвниот притисок. Овие трендови поттикнуваат развој на нови методи кои не само што ја поддржуваат популарната осцилометриска техника, туку обезбедуваат и проценка на крвниот притисок, од сигнали за електрокардиограм (electrocardiogram-ECG, ЕКГ), сигнали од фотоплетизмограм (photoplethysmogram -PPG), комбинација на двата сигнала или додавање на дополнителен параметар за да се добијат попрецизни резултати (Ahmad et al., 2012).

Неинвазивни методи за проценка на крвниот притисок, интензивно се истражуваат во последната деценија. Многу истражувачки групи имаат развиено уреди за набљудување на крвниот притисок кои можат лесно да се носат и на пациентите им овозможуваат континуирано следење на крвниот притисок без да имаат пречки во нивните дневни активности (Baek et al., 2009; Gu et al., 2009; Jungsoo Kim et al., 2007; E. McAdams et al., 2011; E. T. McAdams et al., 2010; Silva et al., 2014; Thomas et al., 2014). Во зависност од влезните сигнали користени за пресметување на крвниот притисок, овие методологии можат да се класифицираат во три категории:

- 1) естимација на крвниот притисок од сигнали на ЕКГ и PPG
- 2) естимација на крвниот притисок само од PPG сигнали
- 3) естимација на крвниот притисок само од ЕКГ сигнали.

Електрокардиографија (ЕКГ) е снимање на електричните активности на срцето. Ова е многу важно за скрининг и дијагностицирање на многу болести. Сигналите

за електрокардиограм се електрични сигнали од срцето. ЕКГ е составено од 5 бранови - P, Q, R, S и T (слика 1.2). Овој сигнал може да се мери со електроди од човечкото тело поставени на соодветни места. Сигналите од овие електроди се доведуваат до едноставни електрични кола со засилувачи и аналогно - дигитални конвертори. Сигналот на ЕКГ треба да биде претставен колку што е можно почист и појасен за да може да обезбеди точни одлуки од страна на лекарите (Ahmad et al., 2012; Parák & Havlík, 2011; Übeyli, 2008; Wu et al., 2016). Како електричен сигнал, ЕКГ е подложен на различни видови на шум. Главните извори на оваа бучава се електрични активности на други мускули на телото, поместување заради дишење, слаб контакт со електродите и опремата или бучава од други електронски уреди. Бидејќи ЕКГ е нестационарен сигнал, нормалните филтри не можат да бидат ефективни за да се отстрани шумот од него. Затоа, се користат неколку техники прилагодливи за вакви типови на сигнали. Постојат различни техники достапни за употреба при отстранување на шум на сигналите: дискретни трансформации на бранови (универзални и локални прагови), адаптивно LMS филтрирање, адаптивно RLS филтрирање и филтрирање Savitzky-Golay (AlMahamdy & Riley, 2014).

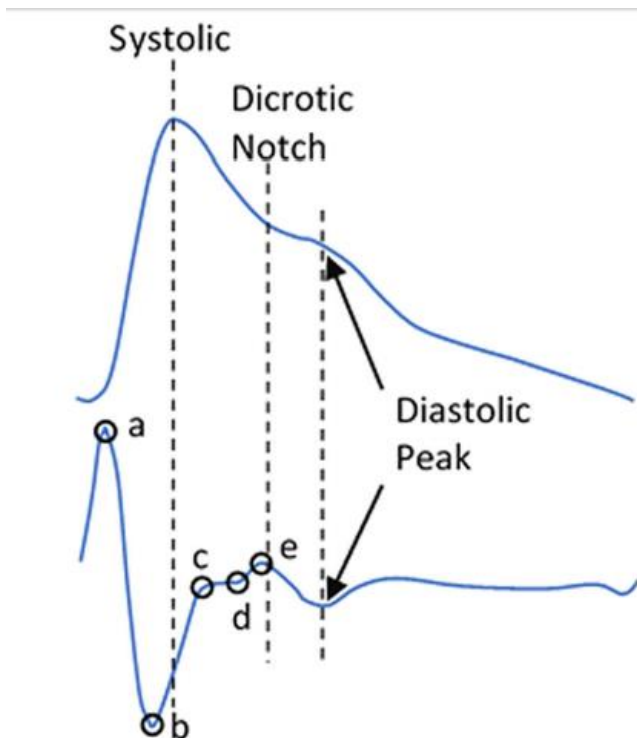


Слика 1.2: Типичен ЕКГ сигнал

Figure1.2 Typical ECG signal

Фотоплетизмографијата (PPG) е оптичка техника што се користи за проценка на протокот на крв во кожата користејќи инфрацрвена светлина. Ова е неинвазивна техника што мери релативни промени во волуменот на крвта во крвните садови близу до кожата. Бидејќи светлината е посилно апсорбирана од крвта во однос на околните ткива, промените во протокот на крвта можат да бидат откриени од сензорите PPG како промени во интензитетот на светлината. Сигналот PPG обично се добива од прстот. Сензорот за прсти PPG се состои од диода што емитува светлина (LED), која честопати е црвена или инфрацрвена, а фотодетекторот PD и LED се на спротивната страна на прстот. Светлината врз кожата се испушта од LED диодата, а фотодетекторот добива мал дел од промените во интензитетот на светлината. Промените се поврзани со протокот на крвта, волуменот на крвта, движењето на сидовите на крвните садови и ориентацијата на црвените крвни зрнца во основното ткиво.

PPG-сигналот (слика 1.3) се состои од АС компонента и DC компонента. Пулсирачкиот дел од сигналот PPG е АС компонента и се добива кога светлината поминува низ артериската крв. АС компонентата е синхрона со отчукувањата на срцето и зависи од промените во пулсирачкиот притисок и пулсирачкиот волумен на крвта. DC компонентата или непулсирачкиот дел е предизвикана од апсорпција на светлина со крв во вените, коските и ткивата (Pilt et al., 2013). Овој сигнал содржи важни информации за варијабилноста на срцевиот ритам, крвниот притисок, дишењето итн. Временскиот период на секој пулс го диктира отчукувањата на срцето и амплитудата со концентрација на разни составни делови за артериската крв и должина на светлината што патува низ артериите (Elgendi et al., 2014) (Joseph et al., 2014; Sangeeta & Laxmi, 2011)



Слика 1.3: Типичен PPG сигнал и бран на неговиот втор извод

Figure 1.3 Typical PPG signal and its Second derivative wave

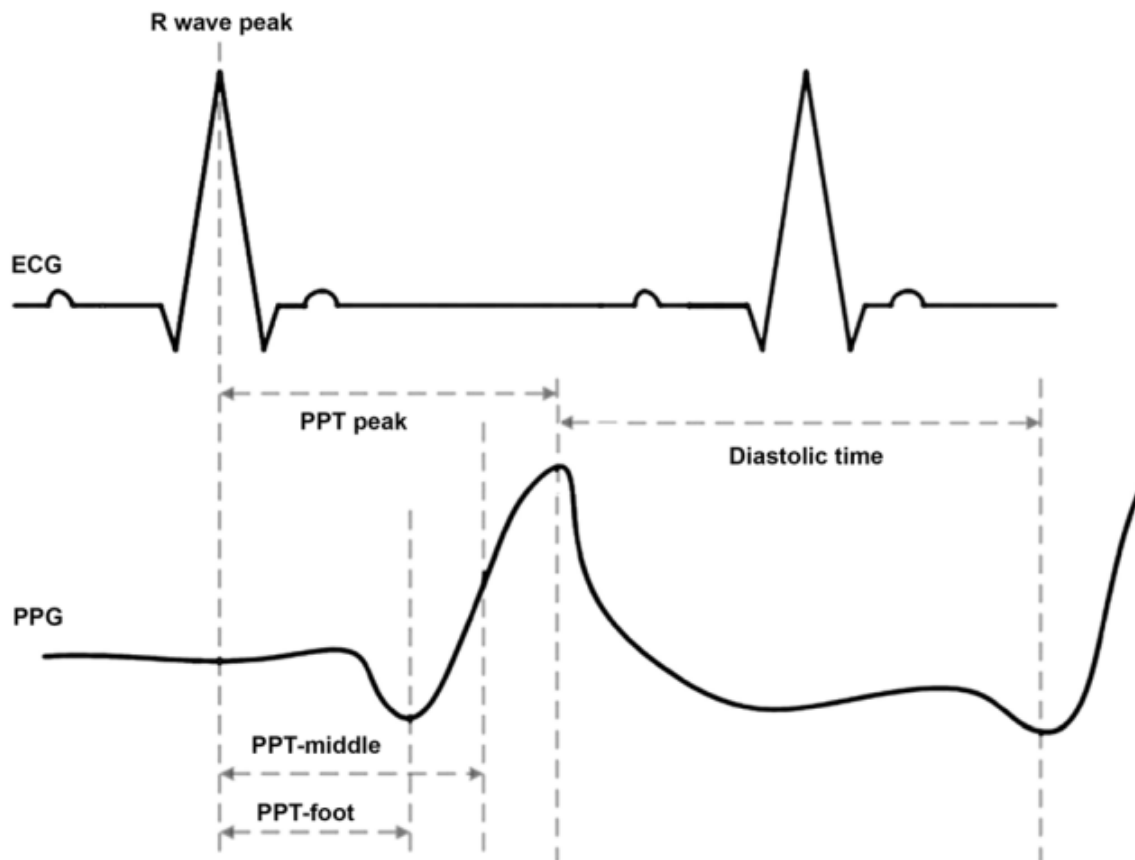
Сигналите на ЕКГ и PPG се тесно поврзани и зависат од возраста на субјектот, неговите дневни активности и неговото здравје, затоа тие можат лесно да се променат, особено сигналот PPG.

1.2.2.1 Естимација на ВР со користење на сигналите ЕКГ и PPG

ЕКГ заедно со PPG сигналите е најкористениот начин за естимација на крвниот притисок кај системите за континуирано следење, особено затоа што се од суштинско значење во пресметувањето на Pulse Transit Time (PTT) (Chandrasekaran et al., 2013; Goli & Jayanthi, 2014; Mohammad Kachuee et al., 2017; Kumar et al., 2014).

Секогаш, кога срцето чука, има наплив на крв од срцето кон сите делови од телото. Брзината на ова движење е директно пропорционална со крвниот притисок. Затоа, потребното време на крвта да патува од срцето до која било специфична локација во човечкото тело е обратно пропорционално со крвниот притисок и тоа одговара на времето за транзиција на пулсот (Pule Transit Time- PTT). Брзината на ова патување одговара на брзината на пулсниот бран (PWV). Со мерење на PTT неинвазивно и континуирано, може да се проценат систолниот и дијастолниот притисок (SBP и DBP) (Gesche et al., 2012; Mohamad Kachuee et al., 2015).

PTT, исто така, може да се опише како потребното време за артерискиот пулсен притисок да патува од аортниот вентил до периферните делови. PTT методот е најчесто користена техника за мерење на крвниот притисок. Постојат два типични методи што се користат за мерење на PTT. Првиот, и почесто користен е да се пресмета разликата помеѓу ЕКГ и PPG, т.е. PTT се дефинира како временска разлика од R-врвот на ЕКГ сигналот до следниот главен врв на соодветниот PPG сигнал (слика 1.4). Обично, ЕКГ се мери од градите, а PPG се мери од показалецот (Y. Chen et al., 2012; Gao et al., 2016; Li et al., 2016)



Слика 1.4: PTT определено со сигналите ЕКГ и PPG

Figure1.4 PTT determined by the EKG and PPG

1.2.2.2 Естимација на BP со користење само на PPG сигнали

РТТ може да се добие со различни PPG сигнали добиени од различни делови на телото, а тоа го прави возможно пресметувањето на крвниот притисок со користење само на PPG сигнали. PPG е едноставен, преносен и низок оптички пулс што може да се користи за откривање на промените во волуменот на крвта. Оваа техника може да се прошири на развој на преносен систем и систем во реално време, способен за постојано следење на пациентите за долг временски период.

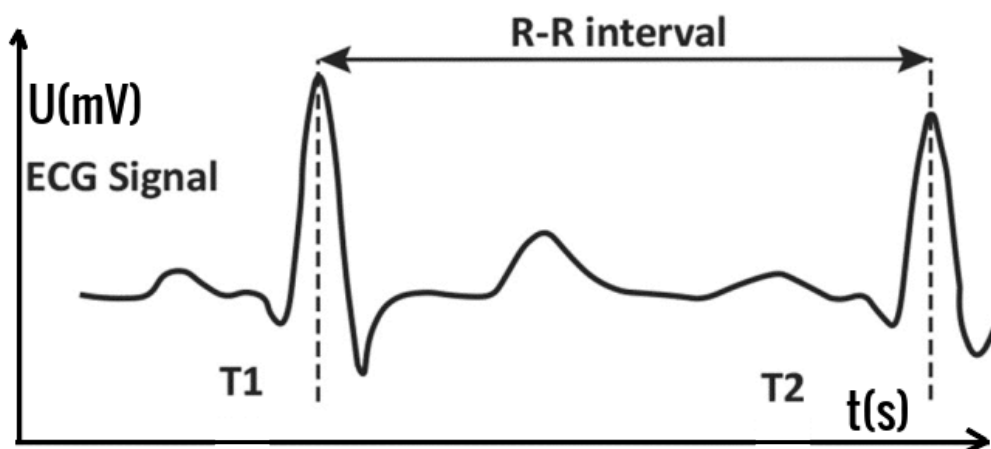
Докажано е дека е можно да се извлечат значајни информации за РТТ од неинвазивни, континуирани форми на мерења на пулсниот бран преку PPG.

Повеќе од алгоритмите и методите за добивање на крвниот притисок од PPG сè уште се во почетната фаза на развој и затоа естимацијата на крвниот притисок на ваков начин не е доволно прецизна.

1.2.2.3 Естимација на BP со користење само на ЕКГ сигнали

Неинвазивно мерење на крвниот притисок, исто така, е можно и со користење само на ЕКГ сигнал (Nemati et al., 2012; Q. Zhang et al., 2017).

Овој метод генерално се заснова на брзината на отчукување на срцето (heartrate - HR). HR може да се определи со делење на интервалот RR од ЕКГ сигналот на 60 (слика 1.5). Овој број е означен како отчукувања во минута (beats per minute – BPM). $HR = 60 / (RR \text{ interval во секунди})$. Нормален опсег на RR интервалот е 0,6 до 1 секунда, со што нормалната вредност на HR би била (60/1 до 60/0.6, односно од 60 до 100 BPM).



Слика 1.5: Приказ на дефиницијата за срцево отчукување преку сигналот ЕКГ

Figure 1.5 An illustration of the definition of Heart Rate (HR)

1.2.3 Уреди за континуирано мерење на крвниот притисок

Бројот на уреди и системи за континуиран мониторинг на крвниот притисок во последната деценија расте рапидно. Системите за набљудување вклучуваат нови технологии на сензори, мобилни технологии, вградени системи, системи што се носат, амбиентална интелигенција, итн., кои се способни да ги следат пациентите во нивните домови додека ги вршат своите дневни активности, без да се мешаат значително во нивниот комфор или во животниот стил (E. T. McAdams et al., 2010; Silva et al., 2014).

Системите за амбиентално помогнато живеење (Ambient Assisted Living System -AALS) (Koceska et al., 2017, 2019; Kulev et al., 2015; Trajkovik, Koceski, et al., 2015; Trajkovik, Vlahu-Gjorgievska, et al., 2015; Vlahu-Gjorgievska et al., 2016) го опишуваат животното опкружување со зголемено присуство на ИКТ технологија, во кое се следат околните услови преку мрежа на сензори и собраните податоци честопати се споени заедно со информациите собрани од системите за набљудување на здравјето и активностите за да овозможи:

1. контрола на животната средина за безбедност и удобност на пациентите,
2. на семејството, пријателите и старателите да им дадете ажурирани информации за статусот на пациентот,
3. информирање за краткорочното и долгорочното управување со здравството и грижата.

Системите за амбиентално помогнато живеење обично се насочени кон постарите лица, но принципите се слични на парадигмата на паметни или интелигентни домови, па затоа може да се применува во низа други релевантни сценарија.

Што се однесува до уредите кои се користат за континуирано мерење на крвниот притисок, можеме да ги класифицираме во три категории.

1. Уреди за мерење на крвниот притисок кои користат паметен телефон.
2. Лесни за мерење на крвниот притисок што се лесни за носење.
3. Систем со сензори и апликации за паметни телефони.

1.3 Цели на истражувањето

Според направените истражувања, од биофизиолошките параметри кои можат континуирано да се следат со помош на сензори, најпредизвикувачки е крвниот притисок, чие континуирано следење сè уште е во фаза на истражување. За негово следење се потребни дополнителни сигнали преку кои може да се процени неговата вредност.

Според разгледаните методологии, сигналите ЕКГ и PPG заедно се најкористените начини за проценка на крвниот притисок користејќи методи на регресија или во последните студии, вклучувајќи невронски мрежи. Но,

заедничкото користење на ЕКГ и PPG сигналите не е многу погодно затоа што се потребни два различни типа на сигнали, а со тоа и повеќе сензори. Со напредокот на системите интегрирани во чип, можат да се интегрираат дури и моќни единици за обработка на графика, а со тоа се нудат можности за обработка на сигнали во реално време и извршување на сложени алгоритми. Ова може да го зголеми истражувањето за континуирана и точна проценка на крвниот притисок директно и само од сигналите на ЕКГ или само од PPG сигналите. Примената на методите за длабоко учење во овој контекст може да се смета и како ветувачки пристап. Според уредите што се користат за континуирано мерење на крвниот притисок, развивање на цели системи составени од сензори кои комуницираат со апликации на паметни телефони, е најсоодветен за употреба во AAL апликациите. Сензорите треба лесно да се носат на ткаенини или на накит, но сепак да бидат во можност да соберат доволно информации за пресметување на крвниот притисок на паметниот телефон.

Цели на нашето истражување се:

- Развивање на систем со иновативна архитектура за континуирано следење на биофизиолошките параметри, посебно на крвниот притисок, за навремено алармирање при некои брзи промени, и преку тоа да се обезбеди навремена и соодветна реакција, избегнување и спречување на некои сериозни ситуации кои можат да бидат со катастрофални исходи.
- За таа цел овој докторски труд предлага архитектура во која податоците се претпроцесираат во мрежните уреди на рабовите од мрежата и на тој начин ја подобруваат ефикасноста и перформансите на самиот систем, истовремено намалувајќи го мрежниот сообраќај и латентноста.
- Системот да биде составен од сензори за собирање на податоци кои лесно можат да се носат и да не пречат во секојдневните рутини.
- Имплементација на алгоритми базирани на вештачка интелигенција, за обработка на сигнали, за да се направи деривација на потребните биофизиолошки параметри во реално време. На тој начин може да се овозможи и брза интервенција кога е потребно.

- Развој на алгоритам за проценка на крвниот притисок, врз база на PPG-сигналот, кој треба да биде прочистен од шумовите добиени од надворешната средина и шумовите добиени со движењето и секојдневното функционирање на човекот. Пресметувањето на вредноста за крвниот притисок, исклучиво базирано на PPG сигнали, е помалку истражуван метод. Од друга страна, PPG-сигналот се добива со лесно носливи сензори кои се многу едноставни и евтини. Оваа техника, за пресметка на крвниот притисок со користење само на PPG-сигналот, може да се користи за развој на преносен ненаметлив систем и систем во реално време, способен за постојано следење на пациентите за долг временски период и во динамични услови.

1.4 Структурираност на трудот

Трудот е структуриран во шест целини и тоа: Вовед, Пресметка во магла, Обработка на PPG-сигналот и отстранување на шум, Методи за машинско учење, Основни карактеристики на PPG-сигналот и Заклучок.

Во првиот дел, Воведот, се дава осврт на биофизиолошките параметри и се нагласува потребата од нивно континуирано следење за одржување на здравјето. Акцентот е ставен на крвниот притисок како значаен параметар, показател на здравствената состојба, чие континуирано следење претставува особен предизвик. Се наведуваат досегашните постигнувања и обиди направени за лесно и ненаметливо, а сепак прецизно и континуирано мерење на крвниот притисок, споменувајќи ги постоечките методологии и уреди дизајнирани во таа насока. На крајот од овој дел се дадени целите за ова истражување.

Вториот дел е Пресметка во магла (Fog computing). Во овој дел наведени се основните карактеристики на оваа, релативно нова архитектура, но во исто време се нагласуваат и сите можни предности и недостатоци, кои таа ги наметнува. Се прецизира разликата меѓу оваа архитектура и архитектурата за пресметка во облак и се наведени причините зошто е донесена одлука за користење на ваква архитектура во нашиот систем.

Третиот дел е наменет за обработка на PPG-сигналот и неговото прочистување од шум и движечки артефакти. Во овој дел, најпрво се дефинира проблемот што е цел на нашето истражување, а подоцна се предложени и алгоритми за негово решавање. Во овој дел е презентиран нашиот алгоритам кој успешно се справува со отстранување на шумот од PPG-сигналот, посебно шумот создаден при придвижувањето. На крај од овој дел се презентирани добиените резултати.

Четвртиот дел е делот во кој се преставени методите на машинско учење, а дел од нив подоцна се применети и во нашиот систем за да се добие прецизно предвидување. Овде се дадени методи за кластерирање регресија и класификација, а е даден осврт и кон невронските мрежи и длабокото учење.

Во петтиот дел се презентирани извлечените карактеристиките од PPG-сигналот и нивната анотација, врз кои на крај се применети методите на машинско учење. Со компарација на резултатите определено е кој од применетите методи дава најдобри резултати за предикција на крвниот притисок.

Шестиот дел е заклучокот во кој е даден осврт на основните поенти од истражувањето и потенцирани се идните можни насоки за проширување на истото.

2. Пресметување во магла (FOG computing)

Поради сè поголемата застапеност на паметните уреди (мобилни телефони, таблети, сензори, мотори) поврзани преку интернет на нештата (Internet of Things - IoT), се смета дека нивната употреба до крајот на 2020 ќе достигне над 50 билиони единици (Evans, 2011). Количината на податоци произведена од ваквите уреди, како и податоците од другите поврзани делови како фабрики, домови, болници е, исто така, многу голема и се смета дека до 2020 ќе биде околу 194 ексабајти (10^{18} бајти) годишно (OpenFog Consortium Architecture Working Group, 2017). Затоа справувањето со толкава количина на податоци, како и податоците добиени од социјалните мрежи, претставува еден од најголемите предизвици со кој не можат да се справат традиционалните IoT и архитектури базирани на облак. Причината за тоа се: големите размери на податоци, разновидноста и потребата од брза обработка на податоците, познати како Big Data, како и големата латентност на облак базирани системи. Едно можно решение е децентрализацијата на апликациите за управување и анализа на податоците, во самата мрежа, користејќи дистрибуиран компјутерски модел (Iorga et al., 2018). Fog Computing (пресметување во магла), претставува термин кој за првпат е дефиниран и употребен од Cisco (Evans, 2011). Некогаш се поистоветува со термините како Edge Computing, Mist Computing, Fogging or Cloudlets. Иако постојат разлики во овие термини на повисоко ниво, сепак можат да се сметат за синоними. Терминот Fog Computing - пресметување во магла, значи дека наместо работа со централизирана облак околина, магла системите работат на самите краеве од мрежата. Тоа значи дека дел од интелигенцијата, процесите и ресурсите се поставени на самите рабови на облакот. На тој начин, Fog Computing, претставува еден голем дел од интернет на сè (Internet of Everything - IoE). Пресметувањето во магла е нова технологија која овозможува повеќе придобивки на различни полиња, особено во делот на IoT. Слично како пресметувањето во облак, пресметувањето во магла обезбедува сервис за IoT корисниците, како што се обработка на податоци и нивно складирање. Оваа нова технологија има за цел на обезбеди можности за обработка и складирање на

податоци, на самите fog уреди, наместо да ги праќа нив до облакот. И облакот и маглата овозможуваат складирање, пресметка и мрежни ресурси (Mahmood & Ramachandran, 2018). Целта на технологијата магла во IoT е да ги подобри ефикасноста и перформансите и да ја намали количината на податоци која се праќа на облакот за обработка, анализа и складирање. На тој начин податоците собрани од сензорите се праќаат до крајните (edge) мрежни уреди за обработка и привремено зачувување, наместо да се праќаат на облак и на тој начин се намалува мрежниот сообраќај и латентноста.

2.1. Карактеристики на пресметувањето во магла

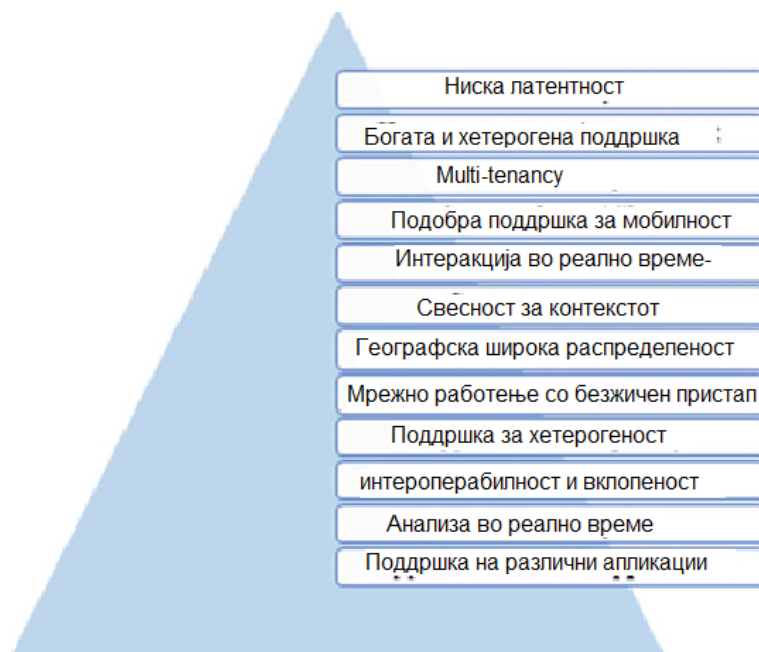
Магла пресметувањето е начин на обезбедување на можности за пресметки и складирање подиректно и поблизу до физичките уреди на некоја организација (што би претставувала рабови (Edge) на облакот (Cloud)). Според OpenFog конзорциумот (Cisco, 2015), пресметувањето во магла претставува хоризонтална архитектура на ниво на систем која го дистрибуира пресметувањето, складирањето, контролата и мрежните функции поблиску до корисникот. Тоа е високовиртуелизирана платформа која обезбедува пресметки, меморија и мрежни услуги помеѓу крајните уреди и традиционалните облак пресметковни податочни центри. Парадигмата на магла обезбедува намалена латентност и свесност за контекстот поради поставеноста на Fog (магла) јазлите, а истовремено подржува вертикално изолирани латентно осетливи апликации со овозможување на сеопфатна, скалабилна и повеќе нивоа мрежна поврзаност.

Магла јазлите обезбедуваат исти видови на услуги како што обезбедува и облак пресметувањето, на пример: софтвер како услуга (Software-as-a-Service - SaaS), платформа како услуга (Platform-as-a Service -PaaS) и инфраструктура како услуга (Infrastructure-as-a-Service -IaaS). Дополнително, архитектурата магла, користи еден или повеќе клиенти како крајни корисници кои меѓусебно соработуваат, или Edge уреди за да се спроведе значителна количина на комуникација, контрола, конфигурација мерење и управување со услуги. Оваа парадигма ги проширува услугите на облак пресметките до рабовите на мрежата. Некои од карактеристиките што ги разликуваат (облак од магла) парадигмите се:

облак околините можат да бидат на големо растојание од организацијата и најчесто не се ни знае каде точно се наоѓаат облак услугите, затоа тие многу зависат од ширината на опсегот на интернетот, додека магла услугите се поблиску поставени до крајните корисници со густа географска распространетост и нудат поголема поддршка за мобилност. Како што е прикажано на слика 2.1, карактеристиките на пресметувањето во магла се:

- Ниска латентност - поради блискоста на fog јазлите до крајните уреди, што резултира со одговор и анализа за побрзо време, односно во помала временска рамка.
- Богата и хетерогена поддршка на крајниот корисник - поради близината на крајните уредите до пресметковните јазли.
- Multi-tenancy во контролирани околии - поради високо виртуелизираната дистрибуирана платформа.
- Подобра поддршка за мобилност - поради повеќе директната комуникација помеѓу Fog апликациите и мобилните уреди.
- Интеракција во реално време - како спротивност од batch процесирањето во случај на облак базирани апликации.
- Свесност за контекстот (Contextawareness) - бидејќи уредите и јазлите во околината имаат знаење и разбирање на самата околина.
- Географска широка распределеност - бидејќи Fog околините се географски дистрибуирани, а тоа игра значајна улога во испораката на висококвалитетни стриминг услуги.
- Мрежно работење со безжичен пристап - тоа е посоодветно за безжични уреди кои бараат временска анализа и комуникација.
- Поддршка за хетерогеност - бидејќи Fog јазлите се со различни форми и се распоредени во најразлични дистрибуирани околии.
- Беспрекорна интероперабилност и подобра вклопеност - за подобра комуникација меѓу уредите од различни производители и низ различни домени.
- Анализа во реално време - која е лесно возможна поради интеграцијата и блискоста на изворите при обработката на податоците.

- Поддршка за разновидни индустриски апликации - поради обработката и анализата во реално време.



Слика 2.1: Карактеристики на пресметувањето во магла

Figure 2.1 Fog Computing characteristic

Овие карактеристики се особено значајни кога пресметувањето во магла се применува во околина како што се: паметни градови, паметни домови, паметни мрежи, паметни медицински апликации или системи кои бараат ургентен одговор. Три основни карактеристики на Fog computing со кои се разликува од Edge и Cloud Computing се: ресурсите, физичката присутност и мобилноста. Пошироката парадигма на магла главно се состои од две други поврзани технологии:

- Cloudlets. Тоа се апликации кои се сместени на рабовите на мрежата, за да одговорат на ниската латентност во машинската комуникација. Cloudlet е јазол на магла, компјутер богат со ресурси или кластер од компјутери (податочен центар), кој е добро поврзан со интернет и достапен за употреба од страна на локализирани мобилни уреди. Магла јазлите можат да бидат или физички или виртуелни елементи кои се цврсто поврзани со крајните паметни уреди или пристапни мрежи. Тие имаат четири главни атрибути, а тоа се: самоуправување, поседување на компјутерска моќност, ниска латентност и основа на облак технологии. Во мрежната архитектура,

Cloudlet се наоѓа на средина во триделната структура: мобилен уред – Cloudlet - Cloud.

- Mobile Edge Computing (MEC). Ова е технологија која е поврзана со мобилната мрежа, преку радиопристапните мрежи (radio access network - RAN) и е во близина на претплатниците на мобилните телефони. Тоа претставува мрежна архитектура која ги овозможува способностите на Cloud Computing и обезбедува околина за ИТ услуги, на работ на мобилната мрежа. Целта е обработката да биде блиска до мобилните корисници, со што се намалува мрежната оптовареност и се зголемуваат перформансите.

Што се однесува до предностите на замаглувањето (Fogging), подетален преглед на моделот на магла сугерира дека станува збор за донесување одлуки што е можно поблиску до податоците и изворите на податоците. Во овој контекст, Hadoop и другите решенија за големи податоци (Big Data) веќе го започнуваат трендот да ја доближат обработката до локацијата на податоците. Fogging е правење на истото само во поголем обем. Постојат очигледни предности кога донесувањето на одлуки е блиску до местото каде се одвива генерирањето на податоци и на тој начин се намалува потребата одредени податоци да бидат обработени во Облак. Ова, пак, од друга страна ги решава и прашањата за безбедност и приватност- бидејќи вредните и чувствителни податоци не треба да патуваат низ Облак мрежите. Некои од предностите на Fogging вклучуваат:

- Густа географска дистрибуција и поддршка за мобилност.
- Ниска латентност, свесност за локација и подобрен квалитет на сервис (QoS).
- Поголема деловна агилност и намалени оперативни трошоци.

Овие предности на концептите и принципите на пресметувањето во магла, треба да се земат предвид при дизајнирање на апликации со некои од клучните принципи за дизајнирање, скалабилност и моделирање и симулација. Ова, исто така, може да помогне да се предвидат карактеристиките на Fog Computing. Сепак, има некои проблеми што треба да се земат предвид кои се слични на проблемите кај Cloud Computing, како што се оние што се однесуваат на безбедноста и приватноста, управувањето со ресурсите, границите на

споделување во случај на дистрибуирани околии и протоколи за комуникација (Naha et al., 2018).

2.2. Прашања и проблеми поврзани со пресметувањето во магла

Иако Fog Computing е нова технологија која ветува многу, сепак важно е да се забележат и разберат различните проблеми што доаѓаат со употребата и примената на ваквиот пристап. Покрај проблемите кои се наследени од обработката во облак (Cloud Computing), постојат и проблеми кои се карактеристични за обработката во магла (Fog Computing).

2.2.1. Безбедност и приватност

Повеќето прашања и проблеми поврзани со безбедноста на дистрибуираните компјутерски околии се однесуваат и на Fog парадигмата. Сепак природата на некои од овие проблеми е различна, бидејќи Fog јазлите се сместени на работ на мрежата, близу до уредите во мрежата. Главните проблеми се однесуваат на автентикација на клиентот на различни нивоа, како и на ниво на мрежни уреди. На пример, ако постои паметен уред кој е поврзан и има IP-адреса. Злонамерниот корисник лесно може да изврши упад во овој уред или да пријави погрешно читање или измама на IP-адреса. Или друг пример, ако постои излезна порта (gateway) која опслужува повеќе Fog уреди, тие лесно можат да бидат компромитирани или заменети со лажни. Исто така, можат да бидат поврзани со злонамерни пристапни точки кои обезбедуваат измамнички SSID. Нападот man-in-the-middle во Fog околина лесно може да се пушти, но тешко може да се адресира. Иако за решавање на ваквите проблеми можат да се користат техники како користење на потпис, автентикација на повеќе нивоа и размена на клучеви со Diffie–Hellman протоколот, сепак тие претставуваат сериозен проблем.

Главните прашања и проблеми може да се групираат во следниве категории:

- Напредни постојани закани (Advance Persistent Threats - APT): Тие се однесуваат на неовластени корисници кои имаат пристап до системи или мрежи и остануваат таму долго време без да бидат откриени.

- Проблеми со контрола на пристап (Access Control Issues - ACI). Овие се однесуваат на неовластени корисници кои успеваат да инсталираат малициозен софтвер за да добијат неовластен пристап за да предизвикаат штета.
- Киднапирање на сметка (Account Hijacking - AH). Овде, целта на нападот е да се добие пристап до кориснички сметки користејќи техники за фишинг со злонамерна цел да го компромитира системот.
- Негирање на услугата (Denial of Service -DoS). Овде, целта на хакерот е да го оневозможи и да го направи недостапен целиот систем или делови од системот што резултира во прекин - како од овластен пристап на корисникот.
- Повреда на податоците (Data Breaches - DB). Ова се однесува на неовластено или нелегално гледање, пристап, бришење, изменување или вадење податоци од страна на индивидуален напаѓач или злонамерна апликација.
- Губење на податоци (Data Loss - DL). Ова се однесува на настан или ситуација која резултира во корумпирање, бришење или правење на недостапност на податоците, од страна на неовластен корисник или злонамерна апликација.
- Злонамерен инсајдер (Malicious Insider -MI) претставува овластен корисник кој ги користи своите дозволи за пристап за штетно, неетичко или нелегално дејствување кое може да предизвика оштетување или оштетување на ресурсите на системот.
- Недоволна длабинска пристојност (Insufficient Due Diligence - IDD) се однесува на недостаток на потребниот стандард на грижа и не исполнување на законската обврска, како резултат на што може да бидат предизвикани штети или загуби.
- Проблеми со делење на технологијата (Shared Technology Issues - STI) се појавуваат кога многу корисници ги споделуваат тие ресурси што можат да резултираат со компромитирање на системот со закани како што се DoS, DL и DB.

2.2.2. Топологија на мрежата Fog и свесност на јазлите за локацијата

Fog мрежите се хетерогени кога управувањето и одржувањето со поврзаноста на голем број разновидни уреди не е лесно. Во тој случај, дизајнот и распоредот на јазли, кои се состојат од различни уреди со разновидни комуникациски протоколи, се едни од клучните проблеми. Архитектурата се заснована на технологијата за виртуелизација, а самото тоа има одредени ограничувања во однос на безбедноста и делењето на граници. Проблемите, во ваков случај се однесуваат на мрежната скалабилност (хоризонтална, вертикална, горе и долу), тополошката распределеност, виртуелизацијата, редундантноста, мерењето на перформанси, следењето, управувањето и оперативните трошоци. Иако проблемите со мрежната топологија можат делумно да се решат со употреба на техники како што е софтверско дефинирано мрежно работење (software-defined networking - SDN) и виртуелизација на мрежна функција (network function virtualisation - NFV), сепак перформансите и скалабилноста на виртуелизираните мрежни уреди претставуваат уште една сериозна загриженост (Yi et al., 2015). Важно е механизмите за интерфејс помеѓу јазлите да бидат динамични, а самите јазли да имаат дополнителен слој кој има вградена амбиентална интелигенција (Aml), па јазлите стануваат свесни за локацијата и контекстот. Но, обезбедувањето на свесноста може лесно да се компромитира поради пропусниот опсег, складирањето и латентноста.

2.2.3. Управување со ресурси

Во секоја мрежа, има проблеми поврзани со откривањето на ресурси, определувањето на крајната точка, алокацијата на ресурси и делењето на ресурси. Fog архитектурата ги има тие проблеми. Најкритичен проблем е дизајнирањето на техники за управување со ресурси кои одредуваат кои модули со апликации за анализа се потребни да бидат доставени до секој работен уред, за да се минимизира латентноста и да се зголеми протокот. За таа цел, потребна е платформа за евалуација која овозможува повторување на квантификацијата на перформансите на полисите за управување со ресурси на IoT или Fog Computing инфраструктурите.

2.2.4. Интероперабилност

Друг клучен предизвик кај пресметувањето во магла е воспоставување на потребното ниво на интероперабилност за да се олесни пристапот до ресурсите. Ова бара пристап во кој различните елементи на Магла инфраструктурата ќе соработуваат. Едно решение за постигнување на интероперабилност е потребата од отворена архитектура која може значајно да ги намали трошоците за развој на Fog апликации, да го зголеми приспособувањето на моделот на обработка во магла и значително да го зголеми квалитетот на услугите на Fog Computing.

2.2.5. Други проблеми

Постојат некои сериозни проблеми како резултат на повеќестанарската природа на околината Магла. Ваквите проблеми се однесуваат на безбедноста и приватноста и договорите за нивото на услуга. Исто така, проблем претставува и недостатокот на соодветни алатки за мерење на поврзаноста, капацитетот, сигурноста, ефективноста и доцнењето. Ваквите проблеми се исти како и кај облак парадигмата заради сличноста на природата на дистрибуираните пресметковни околии.

За да се решат или да се минимизираат проблемите на дистрибуираните пресметковни околии потребни се нови пристапи кои ги задоволуваат најмалку следните побарувања (Mouradian et al., 2018):

- Минимизирање на латентноста.
- Зачувување на пропусниот опсег на мрежата.
- Адресирање на проблемите со безбедност и приватност.
- Безбедно собирање на податоци од различни околии.
- Ефикасно управување со обработка на податоците.

2.3. Пресметка во Облак наспроти пресметка во Магла

Cloud Computing (Пресметка во Облак) е генерален израз за сè што вклучува испорака на хостирани услуги преку интернет (Cearley, 2010). Според (Cearley, 2010), Cloud Computing е стил на пресметка преку која масивно скалабилните IT способности се доставуваат како услуга до надворешните корисници користејќи ги

интернет технологиите. Тоа претставува сеопфатно решение во кое пресметковните ресурси (хардвер, софтвер, мрежно работење, складирање) се обезбедуваат на брз начин до корисниците (Mahmood, 2011). Облак пресметувањето претставува начин на користење на мрежата со далечински услуги, како складирање, управување и обработка на податоци кои се достапни преку интернет, наместо за таа цел да се користи локален сервер или персонален компјутер. Придобивките од користење на обработката во облак се:

- Заштеда на трошоците во однос на капиталните инвестиции - бидејќи организациите можат да ги изнајмуваат или распоредат Cloud ресурсите, без разлика дали тоа се виртуелизиран хардвер, софтвер, меморија или мрежни можности од облак провајдери.
- Намалување на трошоците за ИТ услуги, бидејќи сите услуги поврзани со хардвер, софтвер, меморија и сл., се достапни во облак околината.
- Намалување во одговорностите за управување, а со тоа и фокусот се насочува кон производството и иновациите, бидејќи за одржувањето и распоредувањето на услугите се грижат облак провајдерите.
- Зголемување на деловна агилност за да им се овозможи на претпријатијата да ги задоволат потребите на брзо менување на пазарите - бидејќи најновите технологии можат лесно да се обезбедат од облак провајдерите.

Некои од основните карактеристики на облак парадигмата вклучуваат:

- Самопослужување по барање - овозможување на корисниците да трошат компјутерски ресурси како и кога им е потребно.
- Собирање ресурси - овозможува комбинирање на компјутерски ресурси за да им служи на повеќе потрошувачи.
- Брза еластичност и приспособливост - овозможува функционалностите и ресурсите да бидат брзо и автоматско доделени и скалирани, според предвиденото.
- Определување на мерка за оптимизирање на распределбата на ресурсите - определување на користењето за наплатни цели.
- Проширување на постојниот хардвер и постоечките ресурси - намалување на трошоците за дополнително обезбедување на ресурси.

Услугите засновани на облак можат да бидат од различни видови, но главно се класифицирани како три вида: софтверски услуги (софтвер-како-сервис - SaaS), услуги на платформи (платформа-како-сервис - PaaS) и инфраструктурни услуги (инфраструктура-како-сервис - IaaS). Услугите можат да бидат поделени и на следниве категории: Складирање како услуга, база на податоци како услуга, безбедност како услуга, комуникација како услуга, управување како услуга, интеграција како услуга, тестирање како услуга, деловни процеси како услуга.

Cloud околините можат да бидат од различни видови:

- Јавни или надворешни Cloud околин — Овие се во сопственост, управувани и хостирани од трети страни. Cloud провајдерите ги преземаат одговорностите за инсталирање, управување, обезбедување и одржување. Ваквата разновидност на Cloud овозможува многу поголемо ниво на ефикасност на здружување на ресурсите.
- Приватни или претприемнички Облак околин — Овие се комерцијални мрежи, често центри за податоци, во сопственост и управувани од организациите и со седиште во претпријатијата. На тој начин, претпријатијата можат да имаат поголема контрола на сите аспекти на функционирањето.
- Хибридни Облак околин — Овие се комбинација на приватните и јавните Облак околин, каде одговорностите за управување се поделени делумно на претпријатието и делумно на јавните Облак провајдери. Главната предност кај овие околин е тоа што организациите можат да ги чуваат чувствителните податоци на приватниот облак, а останатиот дел на јавниот.

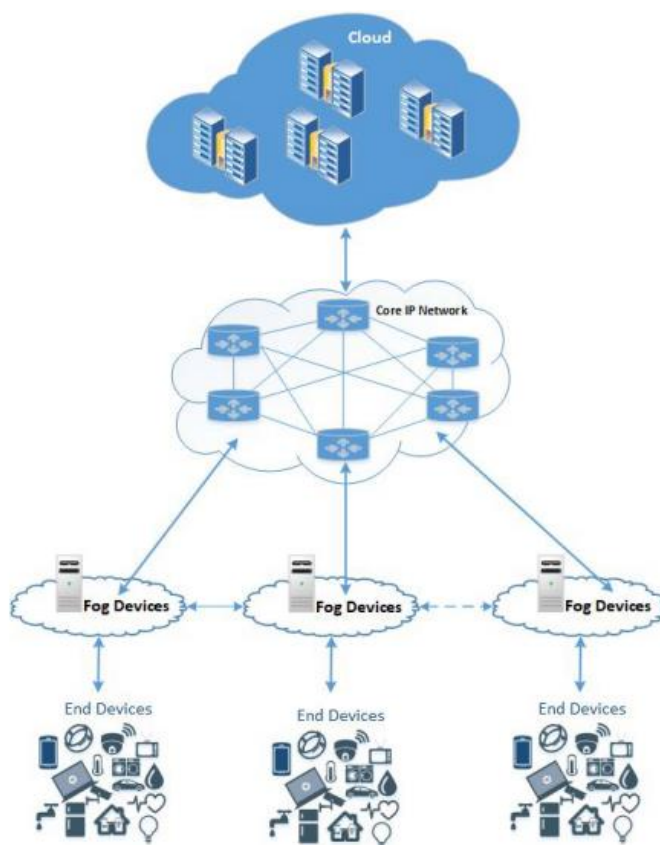
Облак парадигмата нуди бројни придобивки и токму тоа е причината што е доста атрактивен модел, но од друга страна има и бројни проблеми. Некои од проблемите се однесуваат на управување со податоците и управувањето со услуги, следење на производи и процеси, надлежноста на инфраструктурата и достапност на системот, безбедност на информациите и визуелизациите, континуитет на деловно работење, тесни грла на пропусниот опсег и висока латентност, пренесување на податоци низ постојните широкопојасни брзини.

Пресметката во магла е дистрибуирана парадигма што претставува проширување на пресметката во облак до работ на мрежата. Пресметката во магла (Fog Computing) ги олеснува операциите на пресметување, складирање и мрежно поврзување помеѓу крајните уреди и Cloud Computing центрите за податоци. Fog Computing обично вклучува компоненти од апликација кои работат и во облак, и во уредите на рабовите од мрежите. На слика 2.2 е претставен модел на пресметка во магла со 3 нивоа. Во долниот слој се преставени паметни уреди со сензори кои пристапуваат до ресурсите од средниот слој, како и ресурсите директно од облакот. Магла околината од средишниот слој е исто така поврзана и со облак околината во горниот слој. За разлика од тоа, на слика 2.3 е прикажан модел на Fog Computing со 4 нивоа. Табела 2.1 ја илустрира предноста на складирање и обработка на податоци на поблиска локација со користење на Магла, наместо во географски оддалечената Облак околина.



Слика 2.2: Три слоја на Облак Магла моделот (Mahmood & Ramachandran, 2018)

Figure 2.2 Three-layer model of Cloud Fog(Mahmood & Ramachandran, 2018)



Слика 2.3: Модел на пресметка во магла (Naha et al., 2018)

Figure 2.3 Fog computing model (Naha et al., 2018)

Табела 2.1: Предности на Fog Computing во однос на Cloud Computing

Table 2.1 Advantages of Fog Computing over Cloud Computing

	Fog Computing	Cloud Computing
Време на одговор	Милисекунди, дел од секунда	Минути, денови, недели
Период за складирање на податоци	Преоден период	Месеци и години
Апликации	Пр. M2M	Пр. Анализа на податоци
Покриеност на локација	Локално	Глобално

Пресметувањето во Облак бара голем пропусен опсег, а интернетот може да биде несигурен и безжичните мрежи имаат ограничувања. Со употреба на пресметување во Магла, потребниот пропусен опсег е значајно намален. Fog Computing овозможува, пренесените податоци да го заобиколат интернетот,

чувајќи ги што е можно локално на паметните уреди во околината на Магла. Највредните податоци сè уште може да се пренесат преку Cloud мрежите, но голем дел од сообраќајот, особено чувствителните податоци, би можеле да се чуваат надвор од тие мрежи, ослободувајќи го пропусниот опсег на облакот за сите други што го користат. Слично како Cloud Computing, Fog Computing обезбедува меморија, пресметки и апликации до крајните корисници. Но, Fog Computing има поголема близина до крајните корисници и поголема географска дистрибуција (Bonomi et al., 2012b). Во споредба со Облак парадигмата, пресметувањето во Магла ја нагласува близината на крајните корисници и целите на клиентот, обемот на локалните ресурси, латентноста и намалувањето на главниот пропусен опсег за да постигне подобар квалитет на сервис (QoS).

На тој начин, Fog Computing го проширува Cloud моделот до работ на мрежата за да ги опфати апликациите и услугите што не се вклопуваат во Облак парадигмата поради технички и инфраструктурни ограничувања, како што се следниве: Апликации кои бараат помала и предвидлива латентност, географско пошироко дистрибуирани апликации и обработки, побрза мобилност и мобилни апликации, системи за дистрибуирани контроли во големи размери кои бараат побрза обработка. Пресметувањето во магла е погоден за решавање на дел од овие проблеми, на пр. намалување на потребата од голем пропусен опсег со тоа што не се испраќаат сите информации преку Cloud каналите, а наместо тоа, се собираат на одредени пристапни точки. Овој вид на дистрибуирана стратегија, пак, исто така, ги намалува трошоците, ја подобрува ефикасноста, а со тоа и квалитетот на услуга (QoS). Fog Computing претставува уште еден пристап за справување со новиот и популарен концепт на интернет на нештата (IoT). На табела 2.2 се преставени разликите меѓу пресметувањето во Магла и Облак.

Табела 2.2: Разлики меѓу пресметувањето во Магла и пресметувањето на Облак

Table 2.2 Differences between Fog and Cloud computing

	Fog	Cloud
Јазли кои учествуваат	Постојано динамично	Променливо
Управување	Дистрибуирано/Централизирано	Централизирано

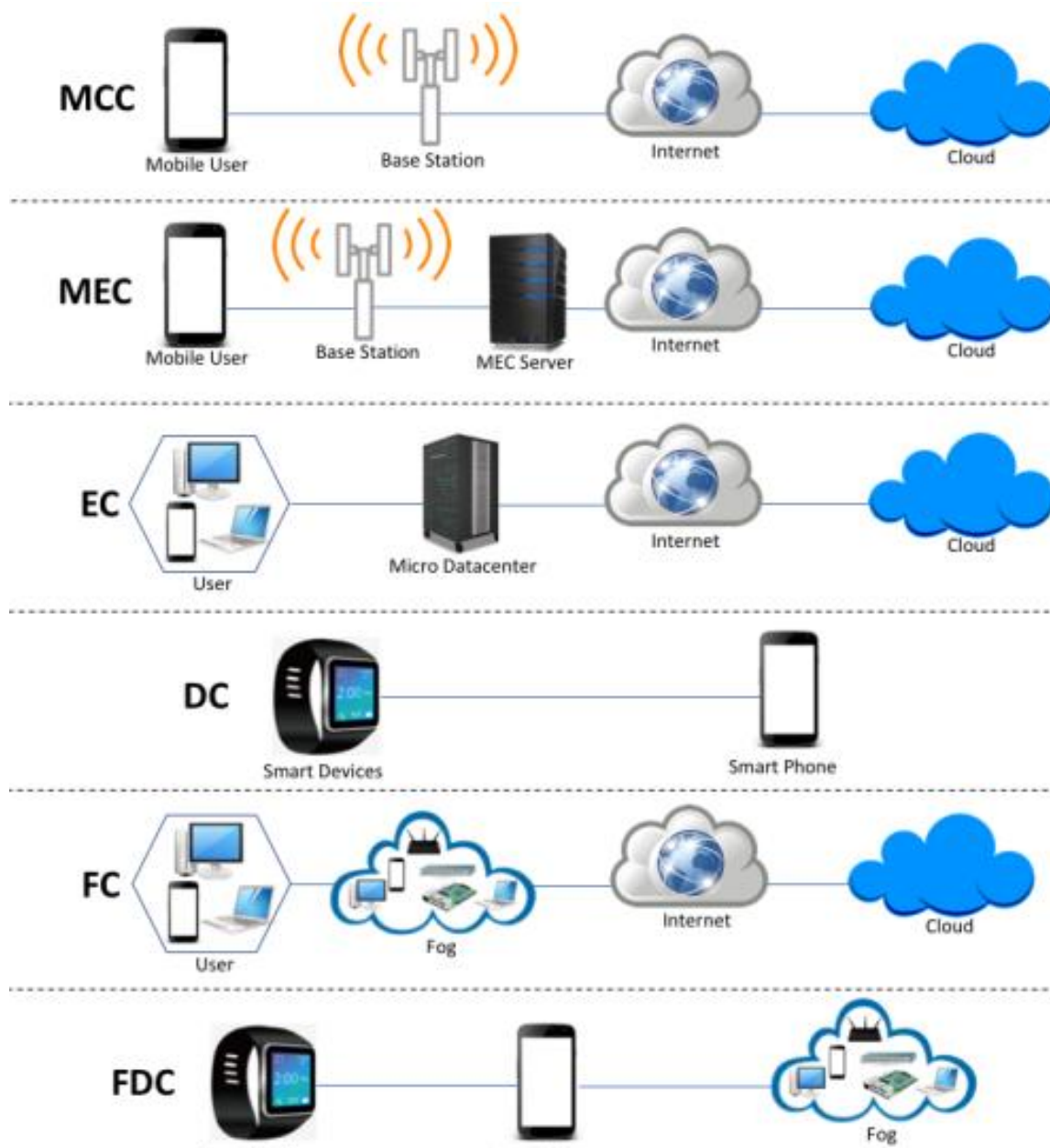
Сметачки уред	Кој било уред со сметачка моќ	Моќен серверски систем
Природа на неуспех	Многу разновидна	Предвидлива
Поврзаност со корисникот	Најмногу безжично	Комбинација на безжична и жична врска со големи брзини
Внатрешна поврзаност	Најмногу безжично	Жична врска
Извор на енергија	Батерија, струја, сончева енергија	Струја
Потрошувачка на енергија	Ниска	Висока
Сметачка способност	Ниска	Висока
Складирачки капацитет	Низок	Висок
Мрежна латентност	Ниска	Висока
Мобилност на јазолот	Висока	Многу ниска
Број на средишни хор-ови	Еден/неколку	Многу
Видови на апликации	Свесни за латентноста	Не свесни за латентноста
Справување со апликации кои работат во реално време	Остварливо	Тешко
Пресметковни трошоци	Ниски	Високи
Трошоци за ладење	Многу ниски	Високи
Време потребно за развој	Многу ниско, може да се инсталира на веќе постојната архитектура	Потребно време за градење на голем магацин

2.4. Поврзани парадигми и технологии

Пресметувањето во Магла користи компјутерски ресурси близу основните мрежи, лоцирани меѓу традиционалните Облак и уредите на работ (edge уредите), за да обезбеди подобра и побрза услуга за обработка на апликациите. Освен Fog Computing, постојат и некои слични компјутерски парадигми и тоа Mobile Cloud Computing (MCC), Mobile-Edge Computing (MEC), Edge Computing, Dew Computing, и Fog-dew computing. При Cloud Computing, сите IoT уреди се директно поврзани со облакот и пресметувањето целосно зависи од облакот. Сепак, сите горенаведени слични технологии не зависат исклучиво од облакот, туку зависат од некои средни уреди за пресметка; некои од нив дури воопшто не бараат врска со облакот (Bonomi et al., 2012 a). На слика 2.4 е прикажана архитектурата на високо ниво на овие технологии.

2.4.1. Mobile cloud computing (MCC)

Далечното извршување на мобилните услуги се врши со поддршка на MCC во близина на крајните корисници. MCC ги надминува пресметковните, енергетските и мемориските ограничувања на ресурси на паметни мобилни уреди. За таа цел се користи лесен сервер за облак (cloudlet) кој е поставен на работ на мрежата. MCC е мобилна компјутерска технологија, која обезбедува неограничена функционалност, мобилност и можност за складирање преку хетерогена мрежна поврзаност. Оваа технологија, исто така, обезбедува унифицирани прилагодливи компјутерски ресурси следејќи го моделот на плаќање по употреба. Исто така, на мобилните корисници им обезбедува пристап до податоци, апликации и облак преку интернет. Се очекува дека оваа технологија во иднина ќе се применува во образованието, урбаниот и руралниот развој, во здравството и во социјални мрежи (Bahl et al., 2012).



Слика 2.4: Архитектура на поврзаните технологии со Fog Computing (Naha et al., 2018)

Figure 2.4 Architecture of related Fog Computing technologies (Naha et al., 2018)

Денес се користат многу апликации кои бараат големи пресметки, како што се проширена реалност, компјутерски визија и графика, препознавање на говор, машинско учење, планирање и донесување одлуки и апликации за процесирање на природен јазик. Но, само дизајнирање на моќни мобилни уреди нема да ги исполни барањата за овие апликации. Наместо тоа, апликациите имаат потреба

од обработка на работ, како и соработка со облакот при сложени пресметки. Така, мобилните пресметковни побарувања имаат потреба од промени во пресметувањето во Облак, на пример, ниска латентност во средното ниво, програмските модели за да се овозможи непречено изведување на далечинско извршување, основни мобилни облак сервиси како што се услуги за присуство, оптимизација на облак инфраструктура за мобилни апликации, мемориски услуги, и сл. Еден од можните решенија на ваквиот проблем се cloudlet-и базирани на виртуелни машини и поставени близу до мобилните уреди (Sanaei et al., 2014) (Satyanarayanan et al., 2013).

2.4.2. Mobile edge computing (MEC)

MEC предлага релокација на пресметковните и мемориските ресурси на базните станици кај мобилните мрежи. MEC може да биде или поврзан или исклучен од податочните центри во облакот. Оттука, MEC поддржува дво- или трислојна хиерархиска структура на апликации заедно со крајните мобилни уреди. Во системот MEC, се воведува нов уред наречен MEC сервер кој треба да биде поставен во близина на базните станици за да обезбеди можности за обработка и складирање на работ. Ваквата компјутерска парадигма ја сочинуваат четири делови, а тоа се: крајните корисници на мобилните уреди, мрежните оператори, Internet infrastructure provider (InPs) и application service provider (ASP). Крајните корисници на мобилна уреди се главни корисници на системот и имаат побарувања за услуги преку корисничка опрема (user equipment -UE). Мрежните оператори ја управуваат и одржуваат работата на базните станици, мобилната мрежа и MEC серверите. InPs ја одржуваат интернет конекцијата и рутерите. ASP се домаќини на апликациските сервиси во мрежите за доставување на содржини (content delivery networks -CDN) или на центрите за податоци. MEC серверот е способен сам да ги обработува побарувањата од корисникот наместо да ги препраќа на далечински Интернет сервиси. Во случај кога не е можно да се обработи или да се изврши барањето на MEC серверот, тогаш тоа се праќа до оддалечени CDN или податочните центри (Beck et al., 2014).

Според Клас (Klas, 2015), MEC е еволуција на мобилните базни станици, односно нивен природен развој. Истовремено претставува меѓусебен развој на

телекомуникациите и ИТ работењето во мрежа. Преку ваквата компјутерска парадигма можат да се испорачуваат различни услуги корисни во различни апликации како, IoT услуги за локација, проширена реалност, видеоаналитика и локална дистрибуција на содржини. Со користење на MEC можат да се добиваат пристапи до локалните содржини во реално време и со ниска латентност. Сепак, главното ограничување на овој систем е инсталацијата на MEC-серверот, кој е специфично наменет за MEC услугите. Скалирањето претставува уште еден голем проблем, особено со зголемувањето на побарувачката на ресурси со текот на времето.

2.4.3. Edge computing

Edge уредите и Edge серверите ги обезбедуваат пресметковните капацитети кај Edge computing. Во принцип, Edge computing не се поврзува со кои било услуги базирани на облак туку повеќе се насочува на страната на IoT уредите. Пресметување на работ (Edge computing) може да се дефинира како која било мрежа или компјутерски ресурс во близина на патеката помеѓу податочните центри на облакот и податочните извори(Shi et al., 2016). Секој паметен уред или сензор може да има извори на податоци, но edge е различно. На пример, cloudlet-от и микроподаточниот центар е edge помеѓу мобилната апликација и облакот, додека IoT излезната порта (gateway) е работ (edge) помеѓу IoT сензорите и облакот. Слично на тоа, ако cloud апликацијата работи на паметен телефон, тогаш паметниот телефон е раб на апликацијата и облакот. Главната мотивација на edge computing е тоа дека пресметките треба да се прават на поблиска локација до изворите на податоци. Во концептот на edge computing, учесниците не само што користат податоци, туку и создаваат податоци и учествуваат во обработката. Edge уредите покрај тоа што бараат услуги и содржини можат и да извршуваат пресметковни задачи од облакот. Edge јазолот може да складира податоци да прави пресметки да обработува и кешира податоци. Edge уредот е, исто така, способен да дистрибуира барања и да дава услуги до корисниците, наместо облакот. Затоа edge уредите треба да бидат добро дизајнирани за да ги исполнат барањата за приватност, сигурност и да се справат со безбедносните проблеми (Satyanarayanan et al., 2009).

2.4.4. Dew computing (DC)

Во тековната компјутерска хиерархија, Dew Computing (Y. Wang, 2015) се наоѓа на најниско ниво кај cloud и Fog Computing околините (Skala et al., 2015). DC го надминува концептот на услуга, складирање и мрежа и оди до концепт на потплатформа, која пак се заснована на концептот на микросервис за кој пресметковната хиерархија е вертикално распределена. DC пристапот ги користи ресурсите како што се сензорите, таблетите и паметните телефони кои се поврзани на мрежа. За таа цел, DC опфаќа широк спектар на мрежни технологии за мрежно работење базирано на ад хок. Според (Skala et al., 2015), DC е многу покорисен во секојдневниот живот во споредба со Fog Computing. Пресметувањето во магла поддржува апликации базирани на IoT, кои бараат помалку латентност и можност за одговор во реално време и динамична мрежна конфигурација, додека DC е концепт за микроуслуга и затоа не зависи од никаков централизиран уред, сервер или облак. Во (Skala et al., 2015) е даден пример во кој DC може да се вклопи во паметен систем за контрола на сообраќајот, во кој единиците за собирање и обработка на податоците се лоцирани помеѓу сообраќајните знаци. Овие единици создаваат колективен преглед на тековните сообраќајни услови. На таков начин, автомобил со ниска количина на гориво ќе биде известен пред да влезе во голем метеж, или хибридниот автомобил ќе биде информиран за префрлување на конвенционалното гориво пред да се приближи кон метежот.

На ваков начин, автомобилите со помалку гориво ќе бидат спасени од несакани ситуации, а хибридните автомобили би можеле значително да ја намалат густината на издувниот чад. Иако концептот е базиран на микросервиси, обработката се завршува во Магла уредите. Кај концептот за пресметка во Магла, не е од клучно значење апликациите да зависат од облакот или да бараат складирање на резултати во облакот. Од друга страна, доколку се чуваат такви информации за сообраќајот, тоа може помогне во стратешкото донесување одлуки за подобрување на управувањето со сообраќајот. Dew computing е брзорастечка област на истражување, чија главна цел е да се искористи целокупниот потенцијал на облакот и локалните ресурси (Axak et al., 2018).

2.4.5. Fog-Dew computing

Во архитектурата на Fog-Dew computing, IoT уредите немаат потреба од активна интернет врска додека се поврзани со локалниот сервер. Локалниот сервер комуницира со облакот и тој е одговорен за давање услуги на IoT уредите (Ahak et al., 2018). Кај Cloud Computing секогаш има потреба од интернет врска, што е главниот недостаток на облакот. Додека облакот не е во состојба да им служи на корисниците без интернет врска, Fog-Dew computing ги спроведува сервисите без интернет врска. Сепак, постојат некои исклучоци. На пример, апликацијата за навигација, Waze, на корисниците им овозможува да се движат офлајн. Оваа особина исто ја има Google Maps. Во овој случај, информациите за картата за одредена област се преземаат до корисничкиот уред и им се овозможува на корисниците да се движат и ако не се поврзани со интернет. Друг пример е Google Drive и Dropbox, каде што корисниците можат да избришат, креираат и ажурираат документи кога не се поврзани на интернет, а потоа да ги синхронизираат промените откако ќе се приклучат на интернет. Сепак, овие услуги не се целосно офлајн и не се потпираат директно на интернет, но не може целосно да се игнорира интернет-врска. Ситуацијата станува посложена кога еден корисник користи повеќе офлајн уреди и на тоа ако се додаде сложеноста што се јавува при повеќекорисничката околина. Ваквите ситуации може да се ублажат со помош на Fog-Dew computing.

Во парадигмата Fog Computing, IoT уредите се поврзани со облакот преку Fog уредите. Fog уредите поврзани со облакот преку основната мрежа. Fog Computing е комбинација на MEC и MCC, но главната идеја кај сите Fog поврзани парадигми е обработката на податоци да се прави на работ. Овие поврзани парадигми се разликуваат едни од други, во однос на поврзаност со интернет и со облакот. Исто така, количината на обработка што треба да се направи на работ се разликува во зависност од побарувањата на услугите. Покрај тоа, типот на уреди што се користат при пресметувањето и складирањето е исто така друг проблем. Табела 2.3 ги прикажува карактеристиките на горенаведените парадигми, заедно со парадигмата за Fog Computing.

Табела 2.3: Карактеристики на поврзани парадигми со Fog

Table 2.3 Characteristics of related Fog paradigms

Сметачки парадигми и апликации	MCC	MEC	EC	DC	Fog-DC	Fog
	Мобилни апликации на сметачки единици близу до корисникот					
	Мобилни апликации	Мобилни апликации	IoT поврзани апликации	Паметни уреди (Здравствено следење)	Паметни уреди (Здравствено следење)	IoT, мобилни видео апликации, паметни мрежи, процесирање на големи податоци
Поврзаност со Cloud	Да	Да или Не	Да или Не	Не	Не	Да
Видови на корисници	Мобилни	Мобилни	Мобилни/Стационарни	Уреди базирани на паметни сензори	Уреди базирани на паметни сензори	Мобилни/Стационарни
Технологија на виртуелизација	хипервизор / контејнер	хипервизор / контејнер	хипервизор / контејнер	контејнер	контејнер	хипервизор / контејнер
Главен сметачки елемент	Сервер на базната станица	MEC сервер	Микро-податочен центар	Паметен телефон	Паметен телефон	Секој уред кој има сметачки и мрежни способности и меморија

2.5. Fog Computing наспроти Edge computing

Бидејќи Fog Computing и Edge computing се многу слични и често пати се поистоветуваат, а на повисоко ниво претставуваат едно исто, овде одделуваме посебен дел за да ги истакнеме нивните разлики.

Fog Computing и Edge computing често се користат за да означат иста архитектура, но сепак постојат и некои разлики кои ги одвојуваат овие две архитектури. Иако Fog и Edge пресметувањето, и двете се однесуваат на постоење на интелигенција, обработка и складирање на работ на мрежата, т.е. поблиску до изворите на податоци, главната разлика е да се направи точно каде се сместени интелигенцијата и обработката. Fog Computing ја става интелигенцијата долу на локалното мрежно ниво од мрежна архитектура, а обработка на податоци во Fog јазолот. Edge computing ги става интелигенцијата,

обработката и комуникациските способности на Edge портата директно во паметни уреди како програмабилни автоматизирани контролори (programmable automation controllers - PAC). Покрај тоа, Edge computing е постар израз кој му претходи на Fog Computing. Додека Edge computing обично се нарекува локација каде што услугите се инстанцираат, Fog Computing подразбира дистрибуција на ресурси и услуги за комуникација, пресметка и складирање блиску до или на самите уреди и системите или кои се во контрола на крајниот корисник.

Кај Fog Computing, податочната комуникацијата помеѓу уредите за генерирање на податоци и Облак околината се состои од повеќе чекори: Комуникацијата прво е насочена кон влезно-излезните точки на програмабилниот автоматизиран контролор (PAC) што стартува контролна системска програма, која ја изведува автоматизација. Потоа, се испраќа со помош на протокол преку излезната порта при што ги претвора податоците во разбирлив формат како што е HTTP. Податоците потоа се пренесуваат на Fog јазол на локалната мрежа што ја извршува потребната анализа и организира пренос на податоци до Cloud за складирање. На тој начин, во околината на Магла има многу врски и многу потенцијални точки на пропусти и дефекти.

Кај Edge computing, комуникацијата е многу поедноставна и поретки се потенцијалните точки на дефекти. Овде, уредите за генерирање на податоци се физички поврзани со PAC, за брза автоматизација како и за паралелна обработка и анализа на податоците. PAC-овите утврдуваат кои податоци треба да се чуваат локално или да се испраќаат до облакот. Така, освен што се намалува можноста од дефекти, има и заштеда на време а и се постигнува рационализирање на комуникацијата, како и намалување на сложеноста на архитектурата (Greenfield, 2016).

Кај пресметката во магла, постои единствен централизиран уред кој е одговорен за обработка на податоци од различни крајни точки во мрежата. Во Edge архитектурата, податоците IoT се собираат и анализираат директно од поврзаните уреди, така што секој мрежен јазол учествува во обработката. Според (Greenfield, 2016), иако Fogging е повеќе посакуван од сервис провајдерите и компаниите за обработка на податоци, Edge архитектурата ја претпочитаат

компаниите со средна опрема кои поседуваат backbone и радио-мрежи. Табелата 2.4 ја дава споредува меѓу Fog и Edge computing во однос на предностите.

Според (OpenFog Consortium Architecture Working Group, 2017), Fog работи со облак, но Edge се дефинира со исклучување на облак. Fog има хиерархиска и глатка архитектурата со неколку слоеви кои формираат мрежа, додека Edge е често ограничена на одделни јазли кои не формираат мрежа. Fog се однесува и на мрежно поврзување, складирање, контрола и забрзување. Fog Computing има способност за peer-to-peer меѓусебна поврзаност меѓу јазлите, додека Edge ги извршува своите јазли во силоси, со што се бара пренесување на податоците назад преку Облакот за peer-to-peer сообраќајот. Архитектурата на Fog моделот се состои од три главни сегменти:

- IoT уреди: овие се поврзани уреди кои генерираат и пренесуваат големи количини на различни структурирани и полуструктурирани податоци.
- Магла мрежа: која прима податоци во реално време од IoT уредите користејќи повеќе различни протоколи за комуникација и врши анализа во реално време.
- Облак околина: која прима податоци за складирање од Fog јазли и исто така врши анализа за деловна интелигенција.

Архитектурата Edge computing се состои од следниве компоненти (Mahmood & Ramachandran, 2018):

- Edge уреди: овие се поврзаните паметни уреди (сензори, активатори и сл.), кои генерираат, анализираат и преземаат други релевантни активности.
- IoTGateway: кој има одговорност за поврзување на Edge уредите со Cloud околината, се справува со разновидни протоколи и ги складира периферните податоци.
- Cloud околина: која прима податоци од излезната порта (gateway), ги анализира тие податоци и испраќа инструкции назад кон излезната порта.

Табела 2.4: Разлики меѓу Fog Computing и Edge computing

Table 2.4 Differences between Fog Computing and Edge computing

	Fog Computing	Edge computing
Предности	Свесност за локацијата,	Сите јазли учествуваат и

	ниска латентност, QoS- но големи ресурси	на тој начин се намалуваат доцнењата
	Податоци блиску до корисникот – но не и за системи кои имаат потреба од ограничени податоци	Локална анализа во реално време
	Вклопување на дистрибуираните податоци со Cloud сервиси	Мали оперативни трошоци и мрежен сообраќај
		Подобри перформанси

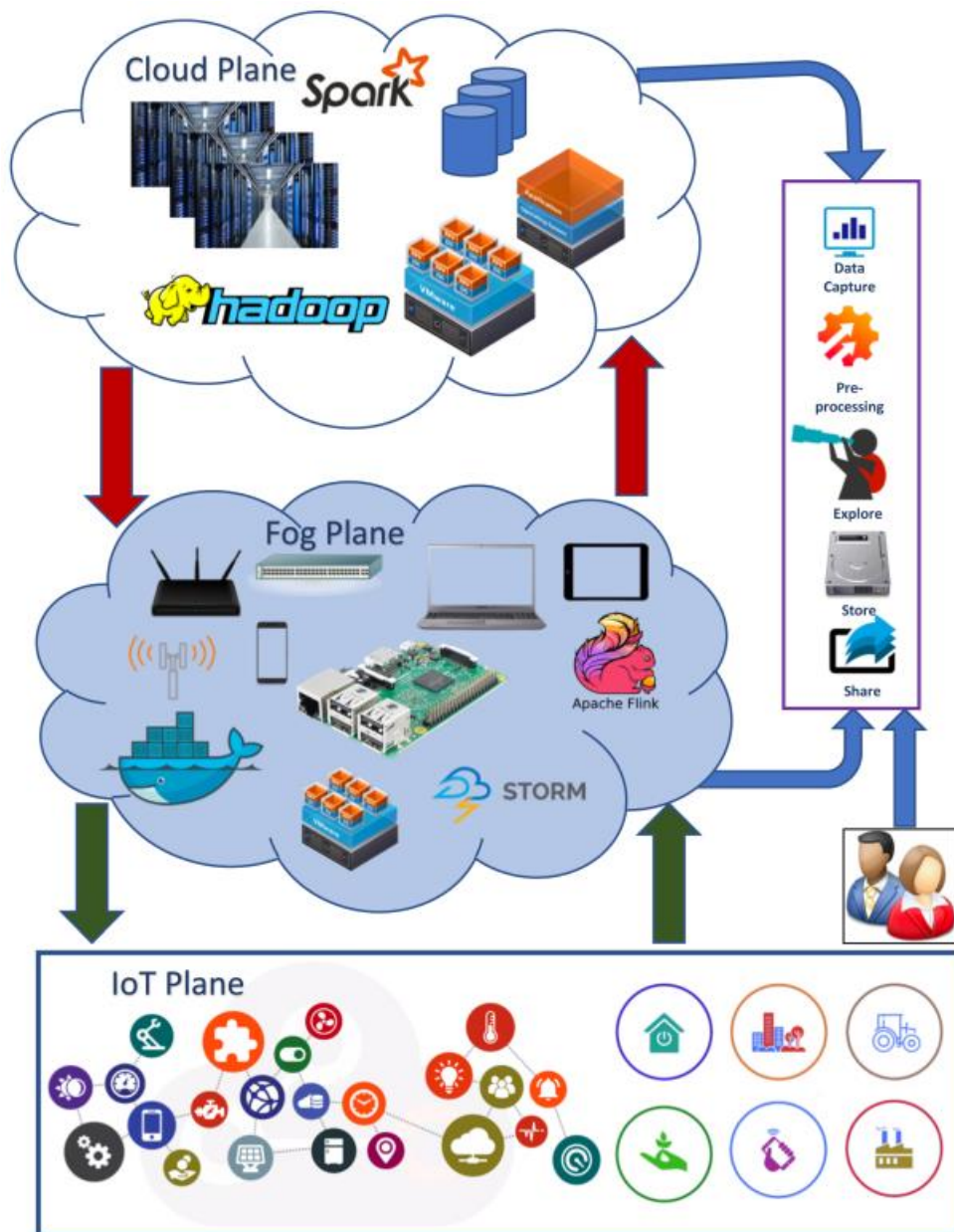
2.6. Архитектура на пресметувањето во магла

За усвојување и распределување на пазарот, Fog Computing мора да имаат стандардна архитектура. До денес нема достапна стандардна архитектура. Сепак во многу истражувачки трудови е презентирана Fog Computing архитектура. Овде ќе дискутираме за високото ниво на Fog Computing архитектурата. Освен тоа ќе сумираме некои од предложените Fog Computing архитектури. На крај ќе претставиме детална Fog Computing архитектура со сеопфатен опис на секоја од компонентите.

2.6.1. Високо ниво на Fog Computing архитектурата

Парадигмата на високото ниво на Fog Computing архитектурата има три различни слоја (слика 2.5). Најважен слој е Fog рамнина (Fogplane). Овој слој се состои од сите посреднички процесирачки уреди. Традиционалните технологии за виртуелизација слични на облак технологиите може да се користат на ова ниво. Со оглед на достапните ресурси, посоодветно е користење на контејнер базирана виртуелизација. Овој слој акумулира сензорски генерирани податоци од IoT слојот и праќа барања поврзани со активирање после обработка. Иако се чини дека проблемот со огромниот размер на податоците е решен со нивната обработка на генерираните податоци на крајното ниво, но сепак милијардите уреди ќе создадат голем проблем со податоците. Всушност на оваа ниво е можно да се користат

мала и средно голема обработка на податоците. Многу истражувања се направени за обработка на големо размерни податоци на ниво на Fog рамнина.



Слика 2.5: Fog Computing архитектура на високо ниво (Naha et al., 2018)

Figure 2.5 High level architecture of Fog Computing (Naha et al., 2018)

Најдолниот слој е IoT рамнината, која се состои од сите поврзани уреди. Уредите во рамнината извршуваат процеси на засекување и активирање. За временски сензибилни апликации, обработката треба да се изврши исклучително на ниво на Fog рамнината, додека облакот може да изврши и други обработки кои не се

временски сензибилни. Се разбира Магла нивото ќе менаџира што треба да се прати на облакот, а што не треба. Клиентите може да добиваат услуги од Магла и Облак врз основа на нивните барања. Сепак Облак рамнината ќе управува со сложената обработка и складирањето.

2.6.2. Различни предложени архитектури за пресметката во магла

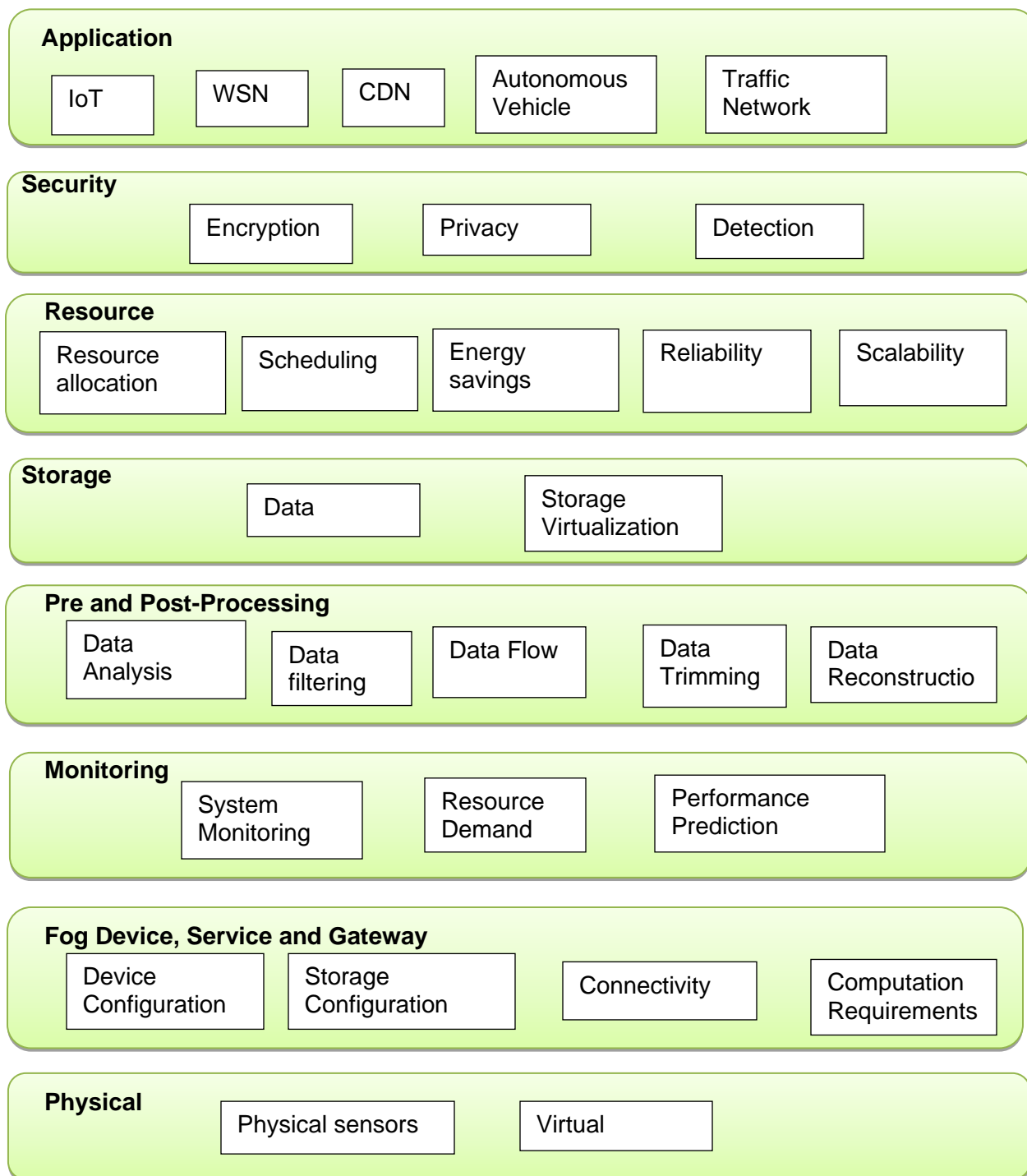
Претставувањето на нивоа е најдобриот начин да се претстави Fog архитектурата. Во многу истражувачки трудови за да се класифицира слоевиот базиран концепт на Магла архитектурата.

Секој има објаснување за своите тврдења. Ако се игнорираат клиентките слоеви, очигледно е дека Магла архитектурата може да се дефинираат три различни слоеви. Како се продолжува со разгледување на повеќе типови на имплементации бројот на слоевите може да варира, така што се зголемува на пет и шест слоеви во Fog Computing слојот.

Aazam и Huh (Aazam & Huh, 2015) имаат преставено шест различни слоеви базирани на конкретни задачи. Од друга страна, (Dastjerdi et al., 2016) дефинираат пет различни слоеви врз основа на мрежната перспектива. Високо ниво на архитектура на Fog Computing е претставено од различни истражувачи каде е вклучена и хиерархијата на Fog архитектурата, OpenFog архитектура, Fog мрежна архитектура, Fog архитектура за Internet of energy, Fog Computing архитектура базирана на нервниот систем и IFCIoT архитектура. Ќе бидат разгледани компонентите на Fog Computing архитектурата, со посебен осврт на OpenFog архитектурата.

2.6.3. Компоненти на Fog Computing архитектура

Fog Computing архитектурата се состои од неколку слоеви. Секој слој и неговите компоненти се прикажани на слика 2.6. Во оваа потсекција дискутираме за различните компоненти на Магла архитектурата. Компонентите се поделени на неколку групи базирани на нивната функционалност, дефинирани како слој. Овие функционалности ќе овозможат IoT уредите да комуницираат со различни Магла уреди, сервери, порти (gateways) и облак. Детален опис на секој слој е даден во продолжение во случајот кога користиме паметен транспорт (smart transportation).



Слика 2.6: Компоненти на архитектура на пресметка во магла

Figure 2.6 Component of Fog Computing architecture

2.6.3.1 Физичко ниво

Основен извор на податоци за Fog Computing се различните форми на податоци емитирани од сензори. Податоците може да бидат генерирани од паметни уреди, сензори за температура, сензори за влажност, паметни домови, CCTV системи за набљудување, системи за мониторинг на сообраќај, само управувачки (self-driving) возилата и така натаму. На пример, ако сакаме да имплементираме паметно менаџирање на сообраќајот, потребно ни е да се следат актуелните сообраќајни состојби од сите патишта и од различни сензори, од уредите покрај патот и од камерите. Сето ова ќе помогне при менаџирање. Исто така, важно е да се предвидат идните сообраќајни побарувања со собирање на податоци од различни GPS сензори. Освен физичките сензори, важна е улогата на виртуелните сензори, доколку настане сообраќајна несреќа, нема да биде возможно да се одлучи само со еден сензор дали патот треба да се блокира или сообраќајот треба да продолжи. Патот може да има една или повеќе коловозни ленти, една лента може да биде засегната додека другата може да овозможи сообраќајот да продолжи, но капацитетот за управување на сообраќајот ќе виде намален поради настанатата состојба. Во овој случај виртуелниот сензор може да помогне при добивање на непосредна состојба за условите на патот, мултиплексирање на сообраќајот и пренасочување на сообраќајот. Поради тоа што физичкиот слој содржи физички и виртуелни сензори, уредите кои генерираат податоци можат да попаднат во некоја од овие групи (Taneja & Davy, 2016).

2.6.3.2 Магла уреди, сервер и порта ниво

Магла уредите, Магла серверите или Магла портите можат да бидат самостојни уреди или IoT уреди (Taneja & Davy, 2016), (Giang et al., 2015). Сепак, очигледно е дека Магла серверот треба да има повисока конфигурација од Магла уред и портата, поради тоа што менаџира неколку Магла уреди. Различни фактори се вклучени за Магла серверите да можат да работат. Ова вклучува улога, хардверска конфигурација, поврзаност, број на уреди кои може да се менаџираат итн. Дали серверот е Магла различен или е дел од IoT уредот се определува од неговата улога. Група на физички и виртуелни сензори се поврзани со Магла уредите. Слично и група од Магла уреди може да биде поврзана на Магла сервер.

Во овој контекст, Магла серверите треба да имаат поголем капацитет за обработка и складирање во споредба со Магла уредите. Специфичен кластер од Магла уреди, кои се поврзани на исти сервер може да комуницираат помеѓу себе ако има потреба. При паметен транспорт (smart transportation) одредени апликациски обработки можат да зависат од други Магла кластери. На пример, доколку некоја апликација треба да пронајде оптимална рута од гледна точка на потрошувачка на гориво, можеби ќе е потребна информација од други сензорски кластери или кластери од Магла уреди. За да се донесе соодветна одлука обработката треба да се изврши во повеќе уреди, Магла уреди и сервери. Магла серверот е одговорен за менаџирање и одржување на информации за хардверска конфигурација, конфигурација на уредите за чување на податоци и конфигурација за поврзување на уредите и серверите. Овој слој, исто така, менаџира со побарувањата за пресметки од различни апликации. Барањата за пресметки зависат од протокот на податоци и тоталниот број на IoT уреди поврзани со Магла уредот, како и од тоталниот број од Магла уреди поврзани на Магла сервер. Комуникацијата помеѓу неколку Магла сервери се одржува во овој слој. На пример, Cisco IOx-800 серија на рутер може да се користи како Магла уред и Cisco Магла податочните сервисни уреди може да се користат како Магла сервер.

2.6.3.3 Ниво за набљудување (Monitoring layer)

Ниво за набљудување секогаш води сметка за перформансите на системот и ресурсите, услугите и одговорите. Компонентите за следење на системот помагаат да се изберат соодветните ресурси за време на работењето. Различни апликации работат во сценарија за паметни системи. Затоа, очигледно е дека може да се појави ситуација кога достапноста на ресурсите за правење на пресметки или складирање на Магла уред нема да биде можна. Слична ситуација може да настане со Магла серверот. За да се справат со вакви ситуации, Магла уредот и серверите бараат помош од другите уреди. На тој начин, компонентата за мониторинг на системот ќе помогне во ефикасно одлучување во ваквите случаи. Компонентата за побарувачка на ресурси ја следи сегашната побарувачка на ресурси и може да ја предвиди идната побарувачка за ресурси врз основа на тековната употреба на ресурсите и активностите на корисниците. На ваков начин,

системот ќе може да се справи со какви било ситуации во кои може да настане и прекин до пристапот на ресурсите. Набљудувачите за предвидување на ефикасноста можат да ги предвидат перформансите на Fog Computing врз основа на оптоварувањето на системот и достапноста на ресурсите. Оваа компонента е потребна за одржување соодветни QoS побарувања кај сервисното ниво. Ако нарушувањето на SLA (Service level agreement) се случува често, тогаш цената на системот за сервисниот провајдер ќе се зголеми заради казната. Иако предвидувањата за перформанси не можат целосно да го отстранат овој проблем, тие сепак можат да го минимизира целокупното нарушување на SLA.

2.6.3.4 Пред и по процесирачко ниво (Pre and post-processing layer)

Ова ниво содржи повеќе компоненти, кои работат на основни и напредни анализи на податоци. На ова ниво, добиените податоци се анализираат и филтрираат, а кога е потребно се прави сечење и реконструкцијата на податоците. По обработувањето на податоците, компонентата за тек на податоци одлучува дали податоците треба да се чуваат локално или треба да бидат испратени до облакот за долгорочно чување.

Главниот предизвик во Fog Computing е да се обработуваат податоците на работ и да се минимизира обемот на податоци што треба да се складираат. Овој феномен се нарекува обработка на поток (stream processing). Во паметните системи, податоците се собираат од многу сензори. Овие податоци се анализираат и распределуваат во реално време за да се добие увид во собраните податоци. Некои од собраните податоци не се употребуваат воопшто. Затоа, во некои случаи, нема потреба да се чуваат сите собрани податоци. Како пример, ако податоците секоја секунда се собираат преку сензори, тогаш можат да се чуваат податоците собрани во рок од една минута или во рок од еден час во зависност од барањата за апликација. На овој начин податоците можат да се скратат и да се зачува огромна количина простор за складирање.

Во друг случај, ако разликата помеѓу вредностите на податоците во одреден временски период не е толку голема, но може да влијае на перформансите, тогаш може да се преземат помалку читања во рок од една минута. На таков начин, може да се филтрира огромната количина на собрани

податоци. Иако точноста не може да биде 100%, побарувањата на апликација можат да бидат исполнети до одреден степен. Реконструкцијата на податоците е една од компонентите на ова ниво. Овој модул се грижи за неисправни и нецелосни податоци собрани од сензорите. Слично на тоа, ако еден или повеќе сензори престанат да работат за време на собирањето, оваа компонента ќе ги реконструира податоците врз основа на моделот за собирање на податоци за да спречи прекин на апликацијата.

2.6.3.5 Ниво на складирање (Storage layer)

Модулот за складирање е одговорен за зачувување на податоци преку виртуелизација на складирањето. Компонентата за резервна копија на податоците обезбедува достапност на податоците и ја ублажува загубата на податоците. Во концептот за виртуелизација на складирањето, базен на уреди за складирање поврзани со мрежа имаат улога на единствен уред за складирање, кој е полесен за управување и одржување. Една од клучните придобивки на виртуелизацијата на складирањето е да се обезбеди функционалност со користење на поевтина меморија и хардвер. На тој начин, нивото на складирање ја олеснува виртуелизацијата на складирањето за да се минимизира комплексноста на системот чување. Кај системот, меморијата може да престане да функционира во кој било момент за време на работата на системот. Затоа, клучно е да се копираат важните податоци за да се ублажат какви било несакани ситуации. Модулот за резервна копија на податоците во ова ниво се грижи за периодични или прилагодени шеми за правење на резервни копии на податоците.

2.6.3.6 Ниво за управување со ресурси (Resource management layer)

Компонентите во овој слој ја одржуваат распределбата на ресурсите и распоредот и се справуваат со проблемите за заштеда на енергија. Компонентата за надежност ја одржува надежноста на распоредот и распределбата на ресурсите. Приспособливоста обезбедува скалабилност на Fog ресурсите во време кога побарувачката на ресурси е голема. Облакот се занимава со хоризонтална скалабилност, додека Fog Computing има за цел да обезбеди и хоризонтална и вертикална скалабилност. Постојат многу дистрибуирани ресурси на системот за мрежно работење, обработка и складирање, поради што се јавува и проблемот

како тие ресурси ќе бидат распределени. Затоа треба да се направи алоцирање, делоцирање и релоцирање на ресурсите, со помош на компонентата за распределување и управување со ресурсите. Друг проблем се јавува поради тоа што многу апликации ќе работат во Fog околината истовремено. За таа цел, потребен е правилен распоред за апликациите.

Посебна компонентата е задолжена за правење на распоред на апликациите, која го прави распоредот според одредени цели и приоритети. Ова ниво, исто така, има компоненти за заштеда на енергија, кои управуваат со ресурсите на ефикасен начин за да се заштеди енергија. Штедењето на енергија, исто така, позитивно влијае на животната средина и помага во минимизирање на оперативните трошоци. Компонентите за надежност се справуваат со побарувањата за надежност на системот врз основа на различни мерки на сигурност. Fog Computing претставува сложен систем што треба да се грижи за сите IoT уреди, Магла уреди и Облакот. Некој уред или врска може лесно да престане да функционира на кое било ниво, така што управувањето со надежноста е важен проблем.

2.6.3.7 Безбедносно ниво (Security layer)

Сите прашања поврзани со безбедноста, како што се шифрирање на комуникациите и безбедно складирање на податоци, се одржуваат од страна на компонентите во овој слој, кои, исто така, ја зачувуваат приватноста на Fog корисниците. Fog Computing е наменет да биде развиен како форма на сметање како што е Cloud Computing. Во концептот за пресметката во Облак, корисникот се поврзува со облакот за услуги, но во концептот за пресметката во Магла корисникот се поврзува со Магла инфраструктурата за услуги додека средниот дел, Fog middleware, управува со одржување на комуникација со облакот. Оттука, корисникот кој има намера да се поврзе со сервис, мора да биде овластен од давателот на услуги (service provider). Затоа, компонентата за автентикација во безбедносниот слој ги обработува барањата за автентикација од корисниците, за да можат да се поврзат со Fog Computing сервисната околина.

За да се одржи безбедноста, клучно е да се одржи шифрирањето помеѓу комуникациите, на тој начин ќе се спречат нарушувања на безбедноста од страна

на надворешни лица. Компонентата за криптирање ги шифрира сите врски од и до IoT уредите и до облакот. Компонентите на Fog Computing главно се поврзани преку безжична врска, така што безбедносните проблеми се клучни. Парадигмата за пресметка во Магла не треба да ги открива информациите за корисниците без нивна согласност. Во денешно време, поголемиот дел од корисниците нормално ја прифаќаат безбедносната политика на сервисните провајдери без да ја читаат. Затоа, треба да се преземе посебно внимание за одржување на приватноста при такви услуги што вклучуваат критични информации поврзани со корисниците.

2.6.3.8 Апликациско ниво (Application layer)

Иако првичната употреба на Fog е за да служи во IoT апликации, неколку други апликации засновани на безжична сензорска мрежа (Wireless sensor network-WSN) и мрежи за испорака на содржини (Content delivery network – CDN), исто така поддржуваат пресметка во Магла. Секоја апликација кај која латентноста е битна, може да ја искористи предноста на Fog Computing. Ова вклучува секаков вид корисни услуги што можат да се вклучат во пресметувањето во Магла и да обезбедат подобар квалитет на услугата и економичност. На пример, апликациите со проширена реалност (augmented reality) треба да го усвојат концептот на Fog Computing заради самата нивна природа. Потребата од обработка на податоци во реално време кај апликациите за проширена реалност може полесно да се задоволи со користење на пресметка во Магла.

2.6.4 Fog Computing референтна архитектура

Дизајнирањето на Fog Computing апликации со ограничени ресурси, побарува добро воспоставени принципи на пресметковни услуги и следење на софтверско инженерски пристап при градењето на апликации со пресметка во Магла што се сигурни и безбедни. Затоа, конзорциумот OpenFog има воспоставено референтна архитектура која обезбедува отворена, дистрибуирана и безбедна платформа за развој на Fog Computing апликации. OpenFog референтната архитектура (OpenFog RA) има за цел да им помогне на бизнис лидерите, софтверските инженери и системските дизајнери во развивањето и одржувањето на хардверот, софтверот и системските елементи неопходни за пресметувањето во Магла.

OpenFog RA е изграден врз множество од осум основни принципи наречени столбови:

- Безбедност: сигурност, доверба.
- Приспособливост (скалабилност): приспособливи перформанси, капацитет, сигурност, безбедност, хардвер, софтвер.
- Отвореност: интероперабилност, комуникација, транспарентност на локација.
- Автономија: на откривање, управување, безбедност, работење, заштеда на трошоците.
- RAS (reliability, availability, serviceability): надежност, достапност, услужливост.
- Агилност: иновација, достапност.
- Хиерархија.
- Програмабилност: приспособена инфраструктура, ефикасно распоредување, ефективно работење, засилена безбедност.

При утврдување на релевантните столбови, се следат стандардите ISO/IEC/IEEE. OpenFog RA нуди голем број на предности што конзорциумот OpenFog ги нарекува SCALE (Security, Cognition, Agility, Latency, Efficiency):

- Безбедност: дополнителна безбедност за да се овозможат сигурни и доверливи трансакции.
- Познавање: свесност за целите на клиентот, за да се овозможи автономија.
- Агилност: брза иновација и прифатливо проширување на веќе постојната инфраструктура.
- Латентност: обработка во реално време.
- Ефективност: динамично здружување на ресурсите од уредите од крајниот корисник што учествуваат во системот.

2.7. Fog Computing апликации

Некои од местата на кои се употребува концептот на пресметка во Магла се:

- Паметни мрежи (smart grids): Fogging обезбедува брза машина до машина (machine-to-machine - M2M) комуникација и човек до машина интеракција

(human-to-machine interactions - HMI), добивајќи поефикасна соработка со Облакот. Магла уредите ги собираат локалните информации и колективно избираат одлука во реално време, врз основа на тоа што се случува во околината.

- Паметни домови и градови: Пресметувањето во Магла овозможува добивање на сензорски податоци на сите нивоа од активностите во домот или во градовите, поврзувајќи ги меѓусебно независните мрежни субјекти, во домовите и градовите, и брзата обработка, за да се создадат повеќе прилагодливи кориснички околина за подобри човечки услови и квалитет на живеење.
- Поврзани возила: Пресметувањето во Магла овозможува идеална архитектура за возило до возило комуникација (vehicle-to-vehicle -V2V), поради близина на уредите вградени во автомобилите, патиштата или пристапните точки. Fogging со свесност за контекстот, ја прави интеракција во реално време помеѓу автомобилите, пристапните точки и семафорите многу побезбедна и поефикасна.
- Автомобили кои сами возат: Овие возила целосно се потпираат на автоматски влез за да може да се изврши навигација. Бавниот одговор во вакви случаи кога возилата се движат со поголема брзина може да бидат опасни, па дури и фатални, така што е потребна брзина на обработка на податоците во реално време и итни одлуки. Fog мрежите се особено погодни за апликации за кои е потребно многу кратко време на одговор помалку од секунда, според Cisco.
- Систем на семафори: Замаглувањето (Fogging) е употреблив при градење на паметни системи за семафор кои ги менуваат сигналите врз основа на дојдовниот сообраќај, за да превенираат несреќи и да ја намалат гужвата. Податоците може да се испратат и до облакот за долгорочни анализи.
- Управување со здравството: пазарот на употреба на обработка во Облак при здравствената нега (Gia, Jiang, et al., 2015) веќе е многу пораснат. Пресметката во облак помага да се забрза тој процес преку зголемување на

поврзаноста на уредите и близината на уредите со пациентите и заедницата на корисници (медицински персонал).

- Медицински уреди за носење: овие уреди сè повеќе се користат од страна на давателите на здравствени услуги за да ги следат промените во состојбите на пациентот, да обезбедат далечинска телемедицина, па дури и да им помагаат на вработените и роботите на лице место, дури и во деликатни процедури како што е операцијата. Овде, надежната и сигурна обработка на податоци во реално време е клучна.
- IoT и сајбер-физички системи (cyber-physical systems - CPS): Fogging има важна улога во CPS при интеграција на физичките и компјутерските елементи на системот, и кај IoT при меѓусебно поврзување на физичките објекти. Нивната комбинацијата веќе го менува светот што се состои од компјутерски-базиран контролирани систем, физичка реалност и инженерски системи.

Други апликации на пресметувањето во Магла се кај: железничката сигурност, реставрација на електрична енергија од паметни мрежи (smart grid), паметни мерачи за паркирање, авто-погони и возови, контрола на воздушниот сообраќај, компјутерска безбедност, IoT сајбер-физички системи, комуникација машина до машина и интеракција помеѓу човекот и компјутерот.

2.8. Иднината на Fog Computing

Атрактивноста на Fog и Edge computing може да резултира во развој на нови деловни модели, со што ќе им помогне на индустриите да растат поефикасно и многу побрзо. Како резултат на тоа, новите индустрии ќе имаат нови понуди и нови архитектури и пристапи за мрежно поврзување. Една област на развој е Fog-as-a-Service (FaaS) во која давателот на Fog услугите распоредува меѓусебно поврзани Fog јазли за да прекрие област со регионални услуги. Ова, од своја страна, ќе обезбеди можности за создавање нови апликации и услуги што не можат лесно да се развиваат со користење на тековните хост и облак базирани платформи. На пример, Магла базираните безбедносни услуги би се справиле со многу предизвици со кои се соочува IoT околината. Појавата на 5G технологии,

развојот на паметни градски апликации и градењето на дистрибуирана компјутерска околина со вградена амбиентална интелигенција, ќе помогнат да се подобри квалитетот на животот и помогнатото живеење преку подобро и поефикасно следење и предвидување на здравјето, со градење на системи за рециклирање и управување со отпад, со поврзување луѓе, паметни алишта, туризам, паметни згради, паметни транспортни и приспособливи паметни домови.

2.9 Примена на пресметувањето во магла кај системите за здравствена нега

Интернет на нештата (IoT) како технологија станува широко прифатена и усвоена во многу аспекти од секојдневен живот на човекот. Оваа технологија овозможува компетентен и структуриран пристап во насока на подобрување на здравјето и здравствената благосостојба на човештвото. Интернет на нештата според очекувањата и предвидувањата има потенцијал да го реконструира здравствениот сектор во однос на социјалните придобивки и економичноста на трошоците (Dohr et al., 2010).

Поради компјутерската природа на IoT, целиот здравствен систем и сите субјекти во него (лица, уреди, лекови) може да се следат и управуваат постојано. Со примена на IoT технологиите во здравството, цената и квалитетот на медицинската нега може значително да се подобри со автоматизирање на задачите кои претходно биле извршувани од страна на луѓето. Во таа смисла, IoT овозможува развој на електронско здравство (e-Health), мобилно здравство (mHealth) и амбиентално потпомогнато живеење кои дозволуваат далечинско набљудување и следење на пациенти кои живеат сами во домашни услови или се лекуваат во болници (Farahani et al., 2018, 2020).

Дизајнирањето на самостојни уреди за носење повеќе не е доволно, наместо тоа, станува неопходно да се создаде целосен систем во кој сензорите поставени на или околу телото на човекот непречено собираат податоци и ги синхронизираат со услугите на облак преку инфраструктурата IoT (B. Xu et al., 2014). Елементите од архитектурата кои најчесто се потребни во здравствените IoT системи се прикажани на слика 2.7.



Слика 2.7: Основна архитектура на IoT систем за здравствена нега

Figure 2.7 General IoT-based health monitoring system architecture

Оваа архитектурата вклучува три главни компоненти:

- 1) мрежа од сензори поставени на и околу телото,
- 2) порти поврзани со интернет,
- 3) поддршка на облак и големи податоци.

Податоците, кои се генерирани од сензорите прикачени на корисниците, им се ставаат на располагање на старателите, членовите на семејството или овластените лица со што им се овозможува тие да можат да ги проверат виталните знаци на субјектот од каде било и во кое било време. Според предвидувањата, сегашните болничко централизиран системи за здравствена нега до крајот на 2020 ќе се развијат во балансирано болнички и домашно централизиран за подоцна во наредната деценија да станат домашно централизиран (Koop et al., 2008). За да може да се реализира таквата еволуција, потребни се нови системски архитектури, технологии и компјутерски парадигми, особено во делот на паметните простори и домените за е-здравство. Промената на парадигмата кон паметни сеприсутни здравствени системи резултира во нови предизвици кои треба да исполнуваат различни побарувања, како што се сигурност, интероперабилност, енергетска ефикасност, мала латентност, мобилност, безбедност и сл. (Rahmani et al., 2018).

Излезните порти се како среден слој помеѓу слојот на сензорите и услугите на облак. Со набљудување на улогата на излезните порти во паметен дом / болница, каде подвижноста и локацијата на корисниците и нештата се ограничени во болничките простории или зградата, може да се забележи дека стационарната природа на портите им овозможува да не бидат со ограничени ресурси или

ограничени во однос на моќта за обработка, потрошувачката на енергија и опсегот на комуникација. Ваквите карактеристики можат да се искористат за засилување на излезните порти со доволна моќ за обработка, интелигенција и мрежна способност, со што стануваат паметни излезни порти за е-здравство.

Сепак, поволните услуги што може потенцијално да ги нуди паметната излезна порта ќе бидат ограничени доколку портата се распореди на самостоен и независен начин. На тој начин, лесно можат да се појават проблемите за приспособливост и подвижност, а со тоа и ефикасноста на решението во таков случај може значително да биде ограничена. Со ова се појавува потребата од воведување на посредник слој на пресметка каде дистрибуираната мрежа на паметни излезни порти обезбедува интелигенција на работ на мрежата и ја олеснува меѓусебната комуникација меѓу слојот на сензорите и слојот на облак.

Ваквата парадигма се нарекува пресметка во магла или пресметка во рабовите на мрежата. Овој начин на работа му овозможува на системот да поддржува непречена подвижност, балансирање на оптоварување, ефикасна приспособливост, мала латентност и развој на апликации со користење на услуги од повеќе сензори и излезни порти.

Во делот на здравствената нега, дизајнирањето ефикасен систем базиран на IoT е предизвикувачка задача заради следниве главни проблеми.

- Прво, избраната технологија за мрежно поврзување на сензорите мора да биде ефикасна во однос на ресурсите и да се прилагоди на апликациите за е-здравство. Јазлите на медицински сензори, особено имплантираните, имаат многу помала моќност за обработка, меморија, брзина на пренесување и снабдување со енергија од сензорите во доменот на другите мрежни сензори.
- Второ, за разлика од вообичаените мрежни сензори каде се користи пренос на податоци на основа на интервали (на пр., следење на температурата и влажноста), апликациите за е-здравство честопати треба да бидат управувани со преноси врз основа на стриминг, при што треба да се земат предвид барањата во реално време. Како резултат на тоа, за време на

процесот на пренесување се троши голема енергија. На пример, преносот на ЕКГ сигналот бара ширина на опсег од 4 kbps по канал.

- Трето, во апликации за повеќе пациенти, како што е случајот во паметни болници, во излезните порти потребни се хардверски платформи со голема моќ за обработка и карактеристики за паралелно процесирање, заради конкурентната природа на работата.
- Четврто, при примена во е-здравство, од најголема важност е и надежноста на системите, затоа неприфатлива е и кратка достапност на системот. Така, како што ќе разговараме во продолжение, ограничените ресурси на јазли на медицински сензори ја прават употребата на портали за општа намена неефикасна во повеќето околности во однос на доцнењето, енергијата и сигурноста.

Постоечките системи за оваа намена користат излезна порта како посредник помеѓу сензорите и складирањето во облак, но и даваат минимална улога на портата, на пример, примена на едноставен пакет правила. Во здравството, особено за далечинско следење на здравјето, се бара високо ниво на сигурност, достапност и стабилност. Покрај тоа, прашањата поврзани со безбедноста и приватноста се од клучно значење. За една излезна порта, во ваков систем за здравство, значајно е да ги задоволува овие специфични барања со прилагодување на портите за потребите во доменот на здравствената нега и обезбедување интелигенцијата да биде поблизу до самите пациенти.

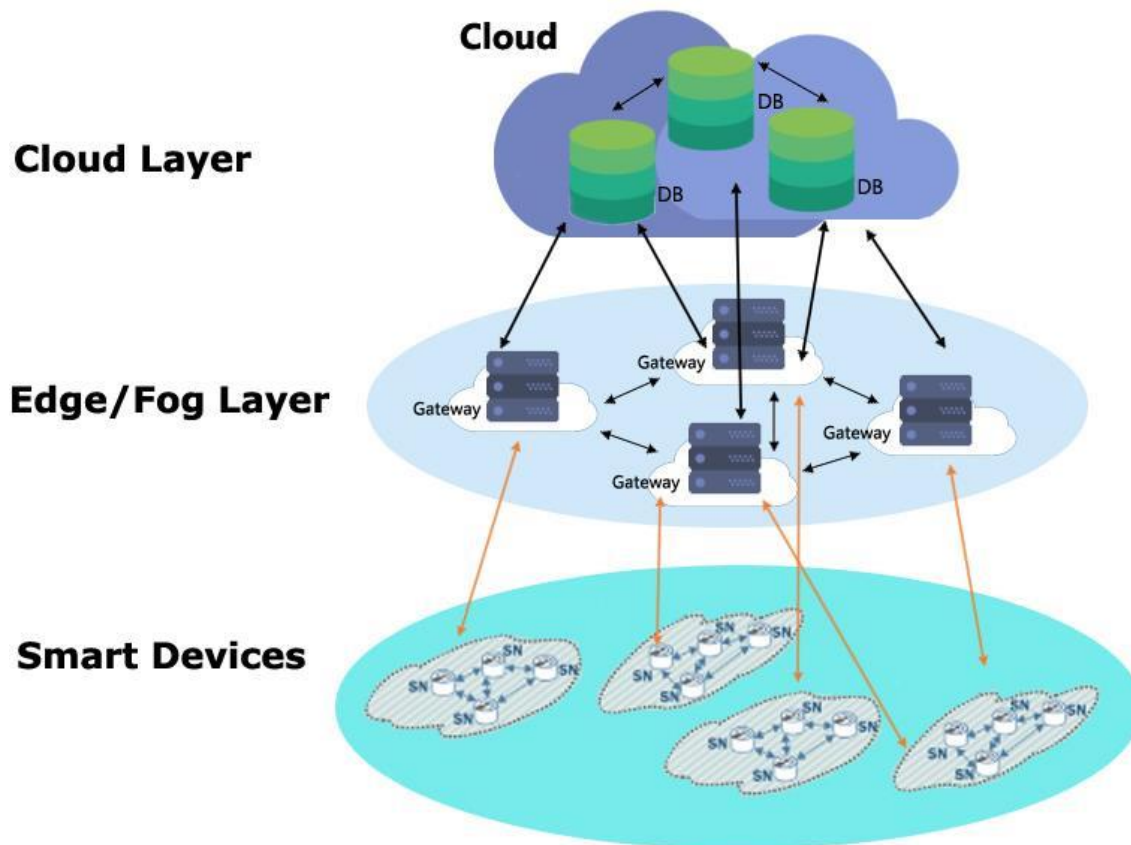
Земајќи го предвид фактот дека во паметна болница или домашна здравствена нега, портата е во непроменета позиција помеѓу WAN/PAN/LAN и мрежата на широк простор (WAN), тоа може да се искористи за собирање на информации за здравјето и околината од ваквите мрежи и соодветно обезбедување на потребните услуги.

Со географско дистрибуирање и вмрежување на паметни порти за е-здравство, може да се формира паметен меѓу слој, за да се обезбедат здравствени услуги кои непречено и ефикасно ќе се извршуваат без да се ограничи мобилноста на пациентите. Основните цели за користење на дистрибуирана мрежа на паметни порти во IoT системите за здравство, се следниве:

1. Да се обезбеди локална обработка на податоци за медицински лица да добијат нотификација во реално време.
2. Да се гарантира безбедност на чувствителните медицински податоци собрани од сензорите и да се зачува приватноста на пациентите.
3. Да се даде интероперабилност за хетерогени платформи и комуникациски протоколи што се користат во мрежните медицински сензори.
4. Да се обезбеди мобилност на пациентите низ делот на покривање од Fog слојот во деловите за болничка или домашна нега.
5. Да се овозможи основната сензорска мрежа да биде поефикасна во однос на енергијата и комуникацискиот пропусен опсег.

2.9.1 Архитектура на системот и улогата на пресметувањето во магла

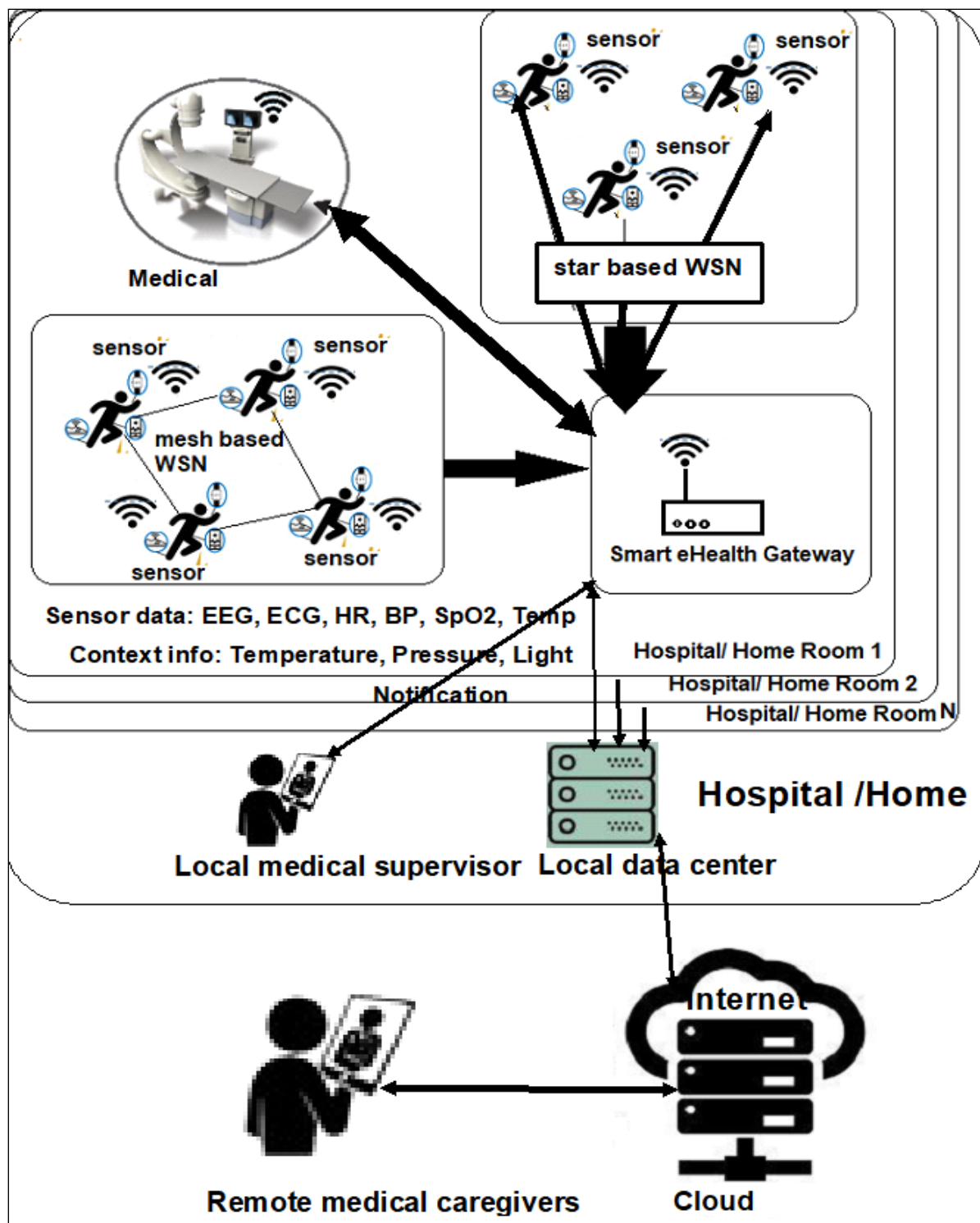
Поради тоа што интернет на нештата се имплементира во големи размери, се очекува многу дополнителни уреди, кои се со ограничени ресурси, да се поврзат на интернет. Поголемиот дел од овие уреди, на пример имплантабилните медицински сензори кои можат да се носат, немаат можност за складирање на податоците кои самите ги генерираат. Наједноставен пристап е да се трансферираат овие податоци во облак за нивно понатамошно процесирање. Земајќи го предвид огромниот број на поврзани уреди, латентноста на конекцијата со облакот би била многу голема. Уште повеќе, овие уреди се со ограничена моќност и ограничен пропусен опсег, па тоа ги прави несоодветни за архитектурата на облак. Пресметките во магла се есенцијална парадигма за поместување кон хиерархиска архитектура на системот. Како што е прикажано на слика 2.8, пресметувањето во магла е интермедијациско ниво меѓу облакот и крајните уреди, со што се надополнуваат предностите на пресметувањата во облак, преку обезбедување на дополнителни сервиси за новите барања во областа на IoT. Целта на паметната излезна порта за е-здравство, како интермедијациски слој, е да обезбеди различни сервиси на рабовите на мрежата.



Слика 2.8: Основна IoT архитектура која се заснова на магла

Figure 2.8 Fog-based IoT architecture

На слика 2.9 е прикажан детален преглед на тоа како компонентите на здравствениот IoT систем можат да бидат организирани на дистрибуиран начин низ трите нивоа, за да бидат користени во паметни болници или паметни домови. Во вакви системи, информациите поврзани со здравјето на пациентот се снимени од страна на имплантирани сензори, или сензори што се носат, со што пациентот е опремен за личен мониторинг на повеќе параметри. Овие податоци за здравјето на пациентот, исто така, можат да бидат дополнети со контекстуални информации (пр. датум, време, локација, температура).



Слика 2.9: Составни компоненти на IoT систем за здравствено следење

Figure 2.9 Components of IoT - based health monitoring system

Свесноста за овие дополнителни информации од опкружувањето овозможува идентификација на невообичаени шаблони и донесување на попрецизни заклучоци во врска со дадената ситуација. Други сензори и активатори (пр. медицинска опрема), исто така можат да бидат поврзани со системите за пренесување на податоци до медицинскиот персонал како што се слики со висока резолуција (пр. CAT скенирање, магнетна резонанца). Системската архитектура ги вклучува следниве главни компоненти:

1. Медицински сензори и мрежа на активатори. Сензорите овозможуваат постојана идентификација и собирање на биомедицинските и контекстуланите сигнали снимени од телото на човекот и просторијата во која тој се наоѓа. Потоа, податоците се пренесуваат до излезните порти преку безжични или жичани комуникациски протоколи како што се Bluetooth или Wi-Fi.
2. Мрежа од паметни излезни порти за е-здравство (e-Health gateways). Ова ниво е изградено од повеќе географско дистрибуирани паметни e-Health порти, на тој начин формирајќи ја маглата. Секоја порта, која поддржува различни комуникациски протоколи, има улога на динамичка допирна точка меѓу мрежата од сензори и локалниот switch. Портата, всушност, ги прима податоците од различни подмрежи, врши конверзија на протоколи и обезбедува други сервиси од повисоко ниво како што се агрегација на податоците, филтрирање и редукција на димензионалноста.
3. Back-end систем. Овој систем се состои од платформа за пресметка во облак која врши имплементација на пренос, овозможува складирање на податоци и нивна анализа. Исто така, обезбедува веб-приказ на клиентот. Собраните податоци во врска со здравјето како и околината претставуваат извор на големи податоци, за статистичко и различно медицинско истражување (пр. откривање на приближување на епидемски болести).

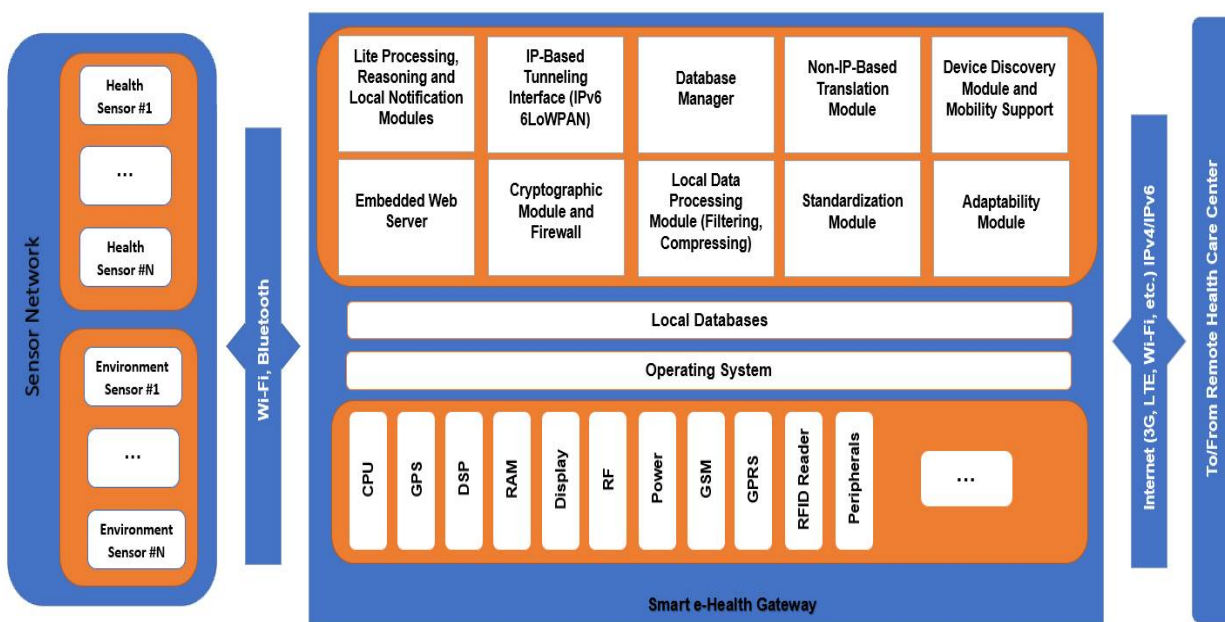
Како што може да се забележи од слика 2.9, интермедијациското ниво е составено од мрежа од паметни порти за е-здравство, поставени на целни стратешки локации, за да понуди повисоко ниво на сервиси и да овозможи подобрување на карактеристиките на системот од повеќе различни аспекти.

2.9.2 Својства и карактеристики на паметни порти за е-здравство во магла нивото

Главната улога на портата е да поддржува разни безжични протоколи и да води сметка за комуникацијата меѓу уредите. Улогата на портата може да се прошири за да овозможи улога на магла со (1) формирање на посебен вид на мрежа на порти и (2) имплементација на неколку карактеристики како што е улога на складиште (т.е. локална база на податоци) за привремено чување на сензорските и корисничките податоци и инкорпорирајќи ги со фузија, агрегација и техники за интерпретација. Овие карактеристики се од суштинско значење за да се обезбеди локална обработка на податоците на сензорите.

2.9.2.1 Локална обработка на податоци

Локалната обработка на податоците е имплементирана како клучна карактеристика на магла компјутерите, за да се обезбедат информации за портата со која податоците локално се анализираат. Според системската архитектура, магла слојот бара континуирано справување со голема количина на сензорски податоци во кратко време и брзо давање на одговор со почитување на различните услови. Ваквата задача станува особено значајна во медицинските итни случаи кои бараат брза реакција (Rahmani et al., 2018). На слика 2.10 е илустрирана архитектурата на паметна е-здравствена излезна порта која користи локална единица за филтрирање, компресија, фузија и анализа на податоци.



Слика 2.10: Паметна архитектура на порта за е-здравство

Figure 2.10 Smart e-Health gateway architecture

2.9.2.2 Филтрирање на податоци

Примањето на податоци од различни сензори го прави неопходно спроведувањето на соодветна претходна обработка на работ од мрежата пред да се изврши понапредна обработка, како што е анализата на податоците. Биофизиолошките сигнали собрани од телото на корисниците се примарни извори на информации за проценка на здравствената состојба на пациентот. Тие обично содржат комплицирани форми со мала амплитуда (во опсегот на милivolти) и разни фреквенции. Кога сензорите се поставени на човечкото тело, шумот се акумулира врз биосигналите и доаѓа до нарушување на квалитетот на сигналот. Шумот се создава од различни извори како што се осцилации на наизменична струја во електричната мрежа, електромагнетно мешање од други електрични уреди и неправилно прицврстување на сензорите на телото на корисниците. Паметната излезна порта за е-здравство во магла нивото може да го реши овој проблем бидејќи директно ги поврзува сензорите. Магла нивото ги прима дигитализираните сигнали од сензорите преку разни протоколи за комуникација.

Иако сензорите може да прават извесно филтрирање за отстранување на дел од шумот уште во самата фазата на собирање на податоци, сепак потребно е посложено филтрирање на податоците во магла нивото.

2.9.2.3 Компресија на податоци

Во податочната комуникација, се користи компресија на податоци за намалување на латентноста на комуникацијата и потрошената енергија за време на трансакцијата. Во доменот на интернет на нештата се користи компресијата и со и без загуба на податоците во зависност од потребите на апликацијата. Компресијата на податоци без загуба има недостаток бидејќи бара прилично тешка пресметка за изведување на комплексни алгоритми. Ваквиот начин на компресија има барања во однос на моќноста и брзината на процесорот. Во системите Health-IoT, во одредени случаи корисни е компресијата со загубите, а во други случаи без загуби на податоци. Сепак, во многу случаи компресијата на податоци со загуба е посоодветна за сензорите со ограничени ресурси, бидејќи има ограничувања, како што е времетраењето на батеријата и достапната моќ за обработка. Сепак, за апликации како што е следењето на ЕКГ во реално време, пожелно е да се има компресија без загуби за да се осигурат сите карактеристики од сите сигнали што се набљудуваат со голема прецизност. Компјутерска магла ја обезбедува потребната компјутерска моќ за ефикасно работење на комплексни алгоритми за компресија на податоци без загуба со исклучување на оптоварувањето од слојот на уредот. На тој начин се овозможува работа во реално време додека се користи компресија на податоци без загуба.

2.9.2.4 Фузија на податоци

Фузијата на податоци му овозможува на системот ефикасно да го намали обемот на податоците, и последователно да ја намали потребната енергија за пренос на податоците. Фузијата на податоци е категоризирана во три класи: комплементарни, конкурентни и кооперативни (Durrant-Whyte, 1988). Комплементарната фузија на податоци може да се изврши во маглата за да се постигне подобро глобално знаење. Пример за ваква фузија е добивањето на температурна разлика помеѓу телото и околината што е обезбедено со два сензорски податока. Конкурентната фузија на податоци може да се искористи на

работ од мрежата на начин преку кој податоците од еден параметар се собираат од различни извори за да се подобри точноста и конзистентноста на резултатите во случај ако некој од сензорите престане да работи. Кооперативната фузија на податоци, исто така, може да работи на работ од мрежата, така што во паметни излезни порти се извлекуваат нови информации од хетерогени податоци собрани од различни извори. На пример, кооперативната фузија на податоци може да обезбеди сеопфатни информации за медицинската состојба на пациентот преку неговите витални сигнали.

2.9.2.5 Анализа на податоци

Чувствителноста на системот се подобрува со примена на локална анализа на податоците на работ од мрежата. Анализата може да му помогне на системот да открие и предвидува вонредни ситуации. На пример, во случај да се исклучат сите извештаи за постарите лица, слојот за магла локално може да понуди обработка поврзана со откривање на податоците, наместо да испраќа параметри на облак и да чека за одговорите. Како резултат на тоа, системот побрзо и посигурно реагира на итни ситуации и имплементира одговори во реално време. Покрај чувствителноста на системот, користењето на анализа на податоци во магла слојот му овозможува на системот да ја минимизира брзината за обработка на критичните параметри.

Понатаму, локалната анализа на податоците и локалните повратни информации од сензорите ја подобруваат сигурноста и конзистентноста на системот во случај на неможност за поврзување на интернет. При долгорочно далечинско набљудување на лица кои страдаат од хронични заболувања, исклучувањето на интернетот може да биде многу веројатно. Во овој случај, магла компјутерите овозможуваат функционалноста на системот да работи локално. Покрај тоа, овозможено е сензорските податоци и резултатите од нивната обработката да се зачуваат локално на магла слојот и подоцна да се синхронизираат со облакот.

2.9.2.6 Адаптивност

Со оглед на примената на магла слојот во различни ситуации, некои предефинирани параметри (на пр., стапка на пренос од сензори на облак) се

поставуваат во зависност од ситуацијата. Неопходно е маглата слојот да биде променлив и прилагодлив со тек на времето, особено во случај на критични настани. Реконфигурацијата може да се прави динамички, користејќи алгоритми за постепено машинско учење, како што се машини со носечки вектори и невронските мрежи.

Преносот на податоци во маглата треба да биде лесно прилагодлив. Тоа се однесува не само на барањата за податоци од сензорите, туку и брзината на пренесување на податоците од маглата до облакот. На пример, при долгорочно следење на пациент кој страда од кардиоваскуларни заболувања, системот треба да научи да ги зголеми стапките на барање (приоритет) за параметрите поврзани со срцето при откривање на абнормален знак поврзан со срцето. Понатаму, брзината на пренесување на облак се врши со доделување на приоритети на различни услуги и параметри. Приоритетот на брзината на пренесување на податоци на облакот за пациенти со акутни заболувања треба да биде повисока, додека пациентите кои страдаат од хронични заболувања бараат помала брзина за пренос. Адаптивното пресметување во маглата ги подобрува перформансите на системот, со зголемување на чувствителноста и специфичноста на критичните параметри.

2.9.2.7 Локално складирање

За да се обезбеди системот непречено да ги обнови податоците, портите треба локално да ги чуваат влезните податоци. Оперативниот систем на портата се справува со локалното складиште и ги чува податоците во случај на нестабилна меморија. Врз основа на видот и значењето, податоците можат да се чуваат локално компресирани или шифрирани. Податоците во складиштето можат да бидат експортирани во формати според медицинските стандарди.

Складирањето на податоците е неопходно за други функционалности на портата. Портата е одговорна за анализа на податоците, компресија, филтрирање и криптирање, и на сите овие функции им е потребно локално привремено складирање. Бидејќи брзината за пренесување на податоците од портата до облакот е ограничена со мрежниот опсег, а пресметките се ограничени со моќта за обработка на портата, во случај на нееднаквост во обработката на податоците и

преносот на податоци, локалното складирање ќе дејствува како кеш за спроведување на континуиран проток на податоци. Складирањето на податоци на порти го прави системот сигурен и стабилен дури и кога мрежата е недостапна. Локалното складиште е управувано од единицата за управување со бази на податоци, прикажана на слика 4.

2.9.2.8 Локално активирање

Во здравствениот систем базиран на IoT, активирањето може да се класифицира во различни форми. Може да се користи во форма на проток на информации, контролирање на медицински активирачи и реконфигурација на сензорската мрежа. Во повеќето од овие случаи, се бара предвидливо и брзо време на реакција. Примери за активирање на брза реакција се прилагодувањето на зачестеноста на електрична стимулација на нервите врз основа на отчукувањата на срцето, или прилагодување на стапката на ослободување на инсулин во автоматските пумпи во зависност од шеќерот во крвта на пациентот и други витални знаци.

Пренесувањето на медицински сигнали на пациентот до контролниот панел за медицинските експерти во реално време, исто така, е чувствителен случај на одложување на преносот, каде што треба да се исполнат минимален број на примероци во секунда. Користејќи ги локалните капацитети за обработка и мрежно поврзување, портата е во состојба да пренесува сигнали како што се ЕКГ и PPG (фотоплетизмограм), до уред за клиент (т.е. таблет), во реално време, без да се потпира на Интернет врска.

Известувањата, исто така, се неопходни карактеристики за паметните порти за е-здравство на работ од мрежата. Системите за здравствено следење честопати треба да ги информираат и предупредуваат медицинските тимови, старателите и пациентот за итни ситуации. Секој неуспех во услугата за нотификација може да предизвика сериозни проблеми и за пациентите и за медицинските третмани. Споредено со облак – сервер, кој може да испраќа известувања преку неколку методи, портата има ограничени ресурси и може да известува само преку некои специфични медиуми. Предноста е тоа што известувањата засновани на магла делуваат независно (на пр., преку локална

мрежа или GSM) дури и за време кога облакот не е достапен, а тоа се прави за да се зголеми довербата на системот и да се осигури дека корисниците можат да добијат критични известувања навреме.

2.9.2.9 Безбедност

Безбедноста може да се смета како едно од најсуштинските барања во здравствените-IoT апликации од причина што небезбедните системи можат да имаат сериозни проблеми. За да се обезбеди високо ниво на безбедност, може да се користат техниките обезбедени на ниво на оперативниот систем како што е IPtable кај Linux. Поточно IPtables и IPFW кои се обезбедени од заштитниот сид на Linux јадрото (Linux kernel firewall) може да се користат за конфигурирање на IP табелата со пакети, што во основа претставува збир од правила на мрежните пакети. Обично IP табелата се конфигурира да дава дозволи на портите за комуникација додека другите порти се блокирани за да се спречи непотребен сообраќај. Пристапната точка или портата (gateway) може да се однесува како вграден веб-сервер при недостапност на мрежата или кога тоа е потребно, може да комуницира прек безбеден HTTPS протокол и да прави автентикација на сензорските јазли за да одржува доверливост, интегритет и автентичност на системот. Покрај тоа што IP табелите имаат определени предности, тие не можат да се сметаат како робусна алатка за безбедност. За да се осигури повисоко ниво на безбедност IP табелите мора да соработуваат со другите напредни безбедносни методи. За решавање на овие проблеми, предложени се различни пристапи (Chakravorty, 2006) (Kambourakis et al., 2007)(Ko et al., 2008). Предложените криптографски операции побаруваат повеќе компјутерски ресурси, а со тоа и енергија така што се неприменливи во овие уреди кои имаат ограничени ресурси. Поради тоа, целта е да се дојде до безбеден и ефикасен пристап за автентикација и авторизација на здравствените -IoT системи кој побарува постоење на определено процесирање на самите рабовите од мрежата. Наместо спроведување на безбедносни методи на самите сензорски јазли кои се со ограничени ресурси, истражувањата се насочуваат кон тоа, извршувањето на тешките и безбедносноповрзаните задачи да се одвива во паметната порта и да

се користи магла пресметката, во ваквите здравствени IoT системи, која е предложена во (Moosavi et al., 2015).

2.9.2.10. Интероперабилност и реконфигурабилност - Interoperability and reconfigurability

Покрај напорите за стандардизација, сепак интероперабилноста игра клучна улога при успехот на здравствените - IoT системи. Со постоењето на хетероген микс од мрежни технологии, протоколи и платформи за имплементирање на IoT-базиран систем, интегрирањето на апликациски слој претставува предизвик. Паметната порта на е здравство игра клучна улога во обезбедувањето на интероперабилноста за различните сензори поврзани со различни интерфејси. Како што е прикажано на слика 2.10, сензорите за биофизиолошките параметри и контекст сензорите се поврзани со паметната здравствена порта (Smart e-Health Gateway) со користење на безжична мрежа или жична конекција и користат различни стандарди (на пр. Bluetooth или Wi-Fi) за комуницирање со портата. Портата се однесува паметно така што лесно ги интегрира овие хетерогени мрежни технологии, протоколи и стандарди со што се овозможува непрекорна размена на информации.

Техничка интероперабилност: Различните системски компоненти во IoT-базираните системи се развиени од различни доставувачи и затоа тие користат различни мрежни интерфејси и стандарди. Во имплементацијата на паметната здравствена порта (smart e-Health gateway), техничката интероперабилност е постигната со директна размена на информациите на самите интерфејси. Адаптационите слоеви на портата ја олеснуваат интерната протокол размена на пораки и формат-конверзијата што се дел од синтаксичката интероперабилност. Не е неопходно портата да ги содржи сите можни интерфејси, туку само тие кои се користат од сензорските јазли. Динамичко вклучување или отстранување на овие интерфејси се управува од страна на карактеристика за реконфигурација на паметната капија.

Синтаксичка интероперабилност: Синтаксичкиот интероперабилен слој се потпира на слојот за техничка интероперабилност. Овој слој се занимава со форматот на пораките и со размената меѓу системите. Откако ќе се достави

пораката примачот треба да ја идентификува содржината на пораката, а со тоа и потребниот протокол за поддршка на модулите во централната порта. Користењето на различните протоколи прави разлики во форматот на пораките. Протокол преведувачите можат да се користат за баферирање на дојдовните пораки и препраќање во друг формат. Овие функции се реализираат со помош на модули, IP базиран тунелирачки интерфејс и не-IP базиран преведувачки модул, прикажани на слика 2.10.

Покрај протоколите на мрежно ниво и медицинските податоци се форматираат во посебен специфичен формат. Податоците во кој било од стандардите на Electronic Health Record (EHR), се реформатираат по потреба од страна на паметната порта. На тој начин, сензорските јазли се ослободуваат од дополнителната обработка за форматирање на податоците според стандардите. А, исто така, се отстранува оптовареноста на комуникацискиот канал, ова се должи на информациите поврзани со стандардот. Испраќањето на неформатирани податоци на портата овозможува сензорските јазли да имаат поголема енергетска ефикасност и пропусен опсег.

Семантичка интероперабилност: за да постои заедничко разбирање за значењето на одредени термини употребени во даден контекст, потребно е вокабуларот прво да биде споделен. Паметната порта за е-здравство е така дизајнирана да овозможи семантичка интероперабилност во двете насоки. Интероперабилноста прво се однесува на другите уреди „заинтересирани“ за собирање на податоци од сензорските јазли. Но, од друга страна портата е поврзана со интернет при што е потребно податоците да се претставуваат во форма разбирлива за човекот.

2.9.2.11. Откривање на уреди и мобилна поддршка

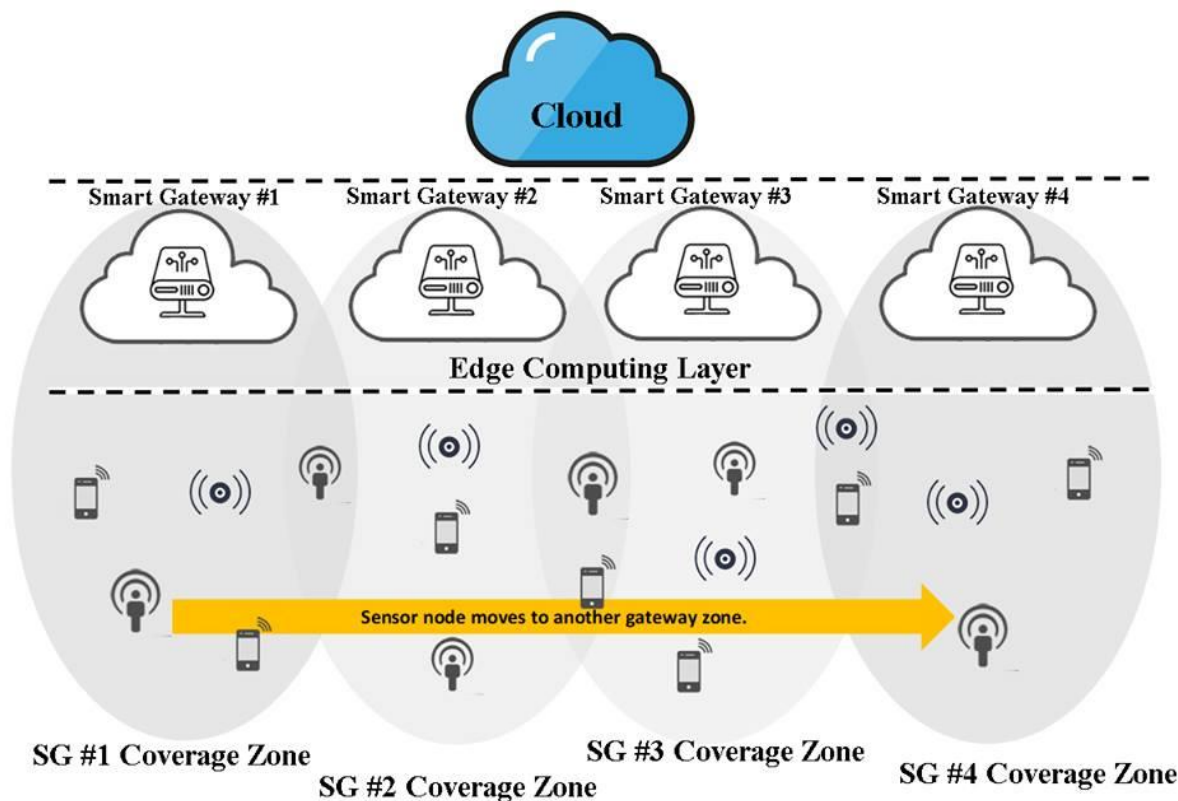
Мобилноста, генерално опфаќа два главни процеса, примопредавање и роаминг, кои се потребни за да се избегне загуба на податоци, прекин на услуги и да се одржи квалитет на услуги (quality of service - QoS). Кај мобилната единица примопредавањето настанува во случај на промена на каналите, додека роаминг настанува кога има движење од една мрежа во друга. Мобилноста, исто така, може да се категоризира во макро и микротипови кои се дефинирани како

мобилност помеѓу различните мрежни домени или во рамките на еден истражен домен.

Постојат неколку методи кои ја подржуваат мобилноста на рутерите на рабовите од мрежата (Rahmani et al., 2018). Во моментот не постои генерален метод кој може целосно да се справи со предизвиците во IoT областа. Исто така, постојат неколку механизми за мобилна поддршка во облак (cloud) или преку дополнителни помошни сервери за далечинско управување. Ова резултира со зголемување на латентноста во процесот на примопредавање, што се должи на големото растојание од мобилниот јазол до облакот. Во околината на здравствената заштита, латентноста на мрежата треба да следи некои стандарди, а тоа само дополнително ја зголемува потребата од дополнителен слој помеѓу јазлите и облакот за да овозможи подобрување на мобилноста.

Поедноставен приказ за тоа како паметните порти се користат во магла слојот за да им помогнат на јазлите за време на поместувањето од една географска локација во друг домен, е прикажан на слика 2.11.

Откривањето на уред помага во идентификувањето на нов јазол кој навлегува во доменот под примарна контрола на поврзаната порта. Преминот на еден јазол од порта #1 до порта #6, е според патека која е прикажана со стрелка (слика 2.11). Секоја порта користи модул со поддршка за откривање на мобилност за да се обезбеди непрекината услуга. Иницијалната конфигурација покажува дека секоја порта управува сет од јазли. Како што јазолот се движи, добива порака која се емитура од страна на портата, која е поврзана со неговиот идентитет. Кога јазолот ќе ја прими вака емитираната (broadcast) порака тој одговара со барање за откривање до соодветната порта. Ова барање е обработено од модулот со поддршка за откривање на мобилност на портата. Во слојот на магла, двете порти (на изворот и на дестинацијата) разменуваат информации во врска со профилот на јазлите и ракуваат со процесот на примопредавање.



Слика 2.11: Мобилност на јазлите во магла пресметката

Figure 2.11 Node mobility in fog computing

2.9.2.12. Енергетска ефикасност на сензорските јазли

Обработката на податоците од страна на сензорските јазли има неколку недостатоци поради ограничените сензорски ресурси. Во некои случаи, и комплексните алгоритми можат успешно да се извршат во сензорските јазли, но за сметка на високи енергетски побарувања. Поради тоа ефикасно решение е процесирањето на сигналите на место во сензорските јазли да се извршува од страна на паметните порти на слојот на магла. Особено ова решение е поприфатливо кога портите не се напојуваат со батерии.

Неколку пристапи се фокусираат за обезбедување на енергетска ефикасност за здравствените-IoT апликации. Во (Otto et al., 2006) се прави процесирање на сигнали во реално време на ниво на сензорските јазли додека во (Gia, Thanigaivelan, et al., 2015) се користи протокол со ниска моќност на трансмисија за да се заштеди енергија при пренос од сензорските јазли. Иако, ваквите техники ја зголемуваат енергетската ефикасност на сензорските јазли,

сепак значителна количина на енергија може да се заштеди преку магла пресметувањето со извршување на задачите од страна на паметните капи.

2.9.2.13 Латентност

За континуиран 24/7 систем за далечинско следење на здравјето, брзото донесување на одлуки и агилните реакции се од суштинско значење за определени акутни болести и други итни случаи, при што времето за процесирање и трансмисија на податоците треба да биде минимизирано. При обработката во облак, каде што податоците се предаваат од сензорските јазли во облакот, мрежните услови се непредвидливи што може да доведе до каснење на одговорот. Ситуацијата е покритична кога е потребно стриминг-базирано процесирање на податоците (на пр. процесирање на ЕКГ или ЕЕГ сигнали). Компаративно, имплементирањето на високоприоритетна податочна анализа во дистрибуираните паметни порти доведува до критични и временски сензитивни одлуки во рамките на локалната мрежа, а тоа го прави системот поробуствен и предвидлив. Потоа, обработените податоци може да се пренесат до облакот за понатамошна анализа. Покрај тоа, во сензорска мрежа од големи размери, локалното процесирање на сигнали во слојот на магла може да го минимизира сообраќајот помеѓу портата и облакот.

3. Обработка и отстранување на шуми движечки артефакти од PPG сигнали

Испитувањето и интерпретацијата на артерискиот пулс за медицински цели било широко применувано уште од антички времиња, кај индиските, египетските, кинеските и грчките цивилизации. Првиот уред наменет за мерење на пулсот бил измислен во 17 век од страна на Sanctorius и бил наречен pulsilogy. Во денешно време, користењето на електрокардиографски (ЕКГ) сигнали е стандарден начин за мерење на срцевото отчукување за медицински цели, мерејќи ја електричната активност во срцето. Со зголемување на интересот за поздрав начин на живот, луѓето имаат тенденција да бидат повнимателни во врска со нивниот пулс и би сакале да имаат неинтрузивен уред за различни активности за мерење на пулсот. Алтернатива за ЕКГ е фотоплетизмографија (Photoplethysmography-PPG). Фотоплетизмографијата се базира на осветлување на крвните садови со употреба на светлечки диоди - лед и мерењето на рефлексивната на светлината во фотодиода.

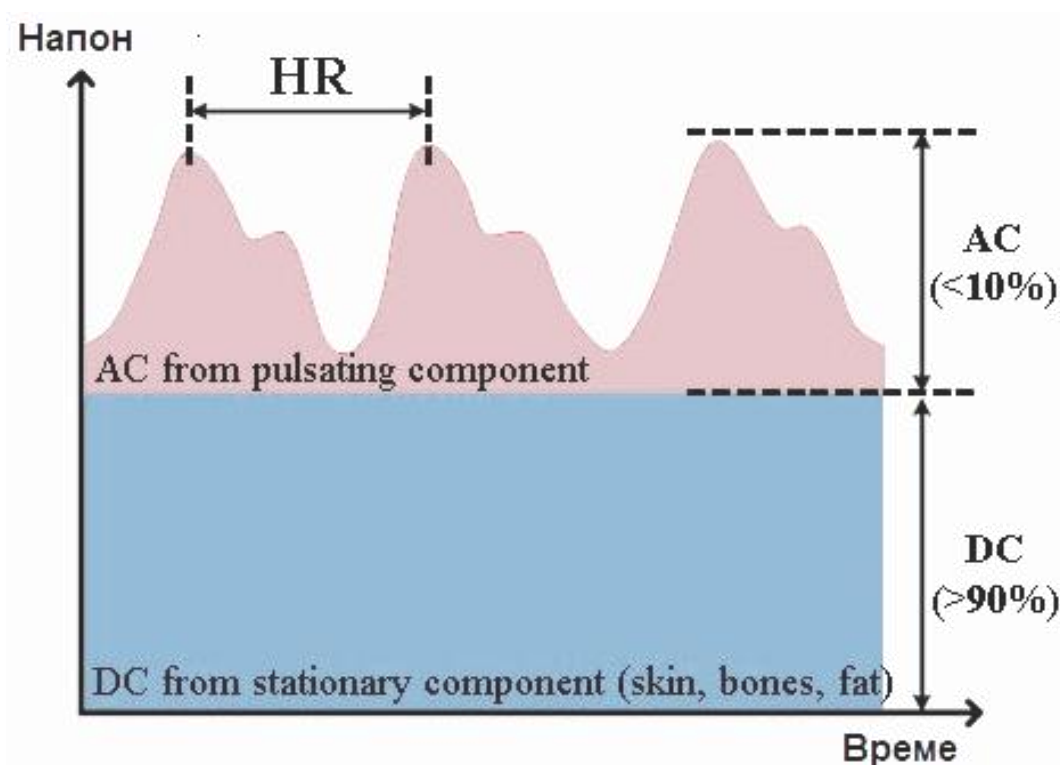
Проблемот со употребата на PPG сензори, поставени на зглобот на раката, е фактот што во тој случај тие се повеќе подложни на движење. При движење, мерењата од PPG сензорите се дисторзирани од таканаречени движечки артефакти (MA). Постојат повеќе техники за процесирање на сигнали од PPG сигнали со MA артефакти, како што се: анализа на независна компонента Independent Component Analysis (ICA) (B. S. Kim & Yoo, 2006), прилагодено отстранување на шум (Adaptive Noise Cancellation (ANC) (Yousefi et al., 2014), TROIKA околина (TROIKA) (Z. Zhang et al., 2015) и Joint Sparse Spectrum Reconstruction (Z. Zhang, 2015). Согласно TROIKA околината, недостатокот на ICA се состои во тоа што отстранувањето на MA, не е на задоволително ниво. ICA претпоставува дека постои статистичка независност на сигналот, меѓутоа ова не е случај за PPG сигналите кои содржат MA. ANC, TROIKA и JOSS во моментот се често користени, а резултатите добиени од TROIKA околината и JOSS изгледаат многу ветувачки. Овие три техники се основа на многу други истражувања и се

оди на нивно подобрување и усовршување како и наоѓање на нови методи кои подобро ќе се справуваат со шумот добиен при движење.

3.1. Фотоплетизмографија

Терминот „plethysmograph“ потекнува од грчкиот јазик и е составен со два збора: „plethysmos“ што значи зголемување, проширување и „graphos“ од графирање, што е збор за пишување. Што означува начин на добивање на траги што покажуваат промени во волуменот на делови од телото. Ова е поврзано со варијации на волуменот како резултат на циркулацијата на крвта во дел од телото што е цел на интерес. Овој термин го вовел Херцман (Hertzman, 1938) во 1938 година за да опише транстектална техника, што ќе овозможи детекција на промените на волуменот на крв при неинвазивна постапка. Принципот се потпира на фактот дека светлината што се пренесува преку живото ткиво, добива промена во времето со брзина што е еднаква со фреквенцијата на чукање на срцето. (Souza et al., 2019)

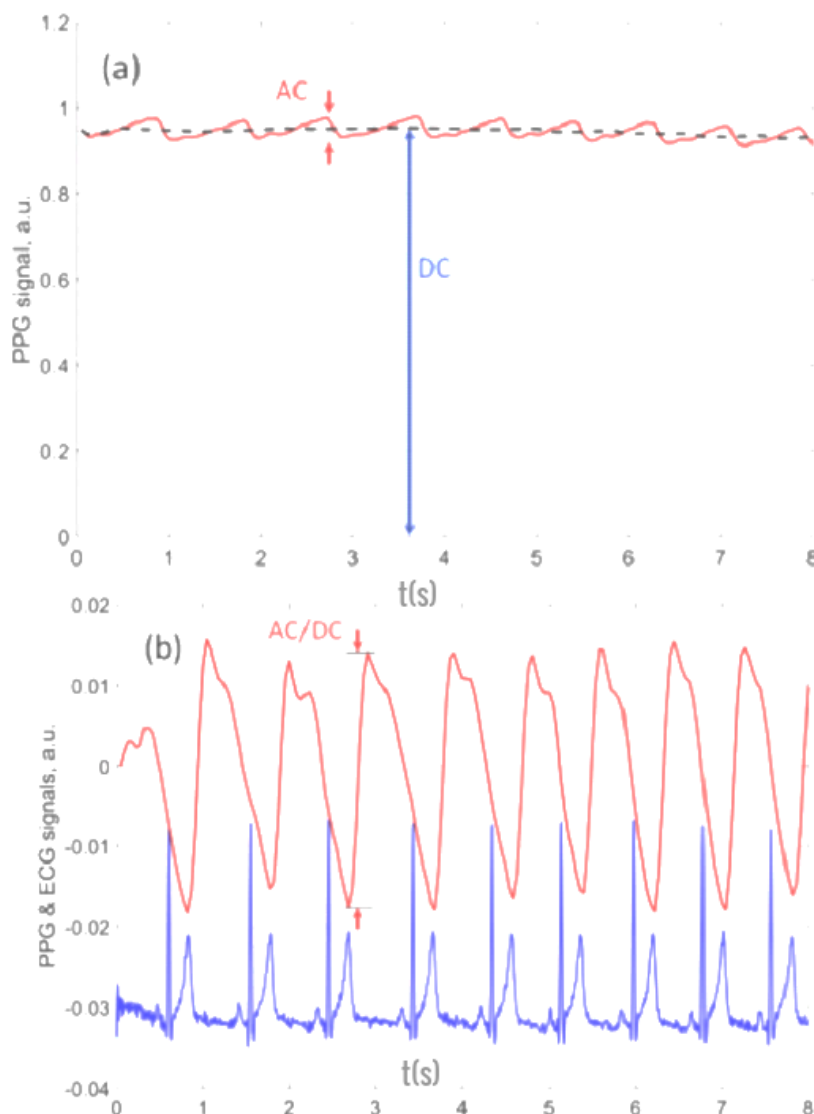
PPG сензорите се неинвазивни, лесно носливи и денеска имаат релативно ниска цена. Тие се користат за постојано следење на промените на волуменот на крвта во капиларните садови користејќи го својството на рефлексција и апсорпција на светлината, произведувајќи употребливи биомедицински информации како што се брзина на отчукување на срцето (heartrate - HR), заситеност на кислород и мерење на фреквенцијата на дишење. Сигналот се состои од две главни компоненти: стабилна DC компонента, која многу бавно се менува и пулсирачка AC компонента (слика 3.1). DC делот одговара на апсорпција на светлината од страна на непулсирачката крв, кожата, коските и други ткива, додека AC компонентата ја претставува апсорпција на светлина со пулсирачката промена на волуменот на крвта помеѓу систолните и дијастолните фази на срцевиот циклус. AC компонентата се води според принципот дека количината на светлина што се апсорбира е обратнопропорционална со пумпната акција на срцето.



Слика 3.1: AC и DC компонента на PPG-сигналот

Figure 3.1 AC and DC component of PPG signal

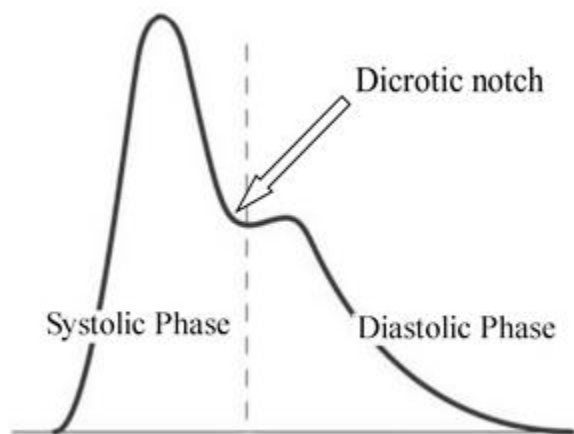
Како што може да се види на слика 3.2 (a.u- arbitrary units, произволни единици), AC компонентата ја преведува интеракцијата помеѓу промените во волуменот на крвта со сензорот. AC компонентата е обезбедена со срцева синхрона варијација во крвта, што се случува заради отчукувањата на срцето. Оваа компонента ги прикажува промените во волуменот на крвта, што ги претставува систолните и дијастолните фази. Систолната и дијастолната фаза се прикажани на слика 3.3.



Слика 3.2: Типична форма на бранот на PPG-сигналот: а) AC компонента (црвена крива, горната крива); DC (црна крива, испрекинатата) е бавно варијабилна компонента; б) сооднос AC/DC, ЕКГ за истиот сигнал е претставен во сина боја(Kamshilin & Margaryants, 2017)

Figure 3.2 Typical form of PPG signal wave: a) AC component (red curve, the upper curve), DC (black, dotted component is a slowly variable component. b) AC/DC ratio.

The ECG for the same signal is present in blue, the bottom curve (Kamshilin & Margaryants, 2017)



Слика 3.3: Систолна и дијастолна фаза кај PPG-сигналот и дикротичен зарез

Figure 3.3 Systolic and diastolic phase in the PPG signal and dicrotic notch

Како што е прикажано на Слика 3.3, првиот врв е патување на крвта напред од срцето до периферните делови наречено систолен врв, а вториот е дијастолан врв, што е рефлектиран бран што се враќа од периферните делови. Аортна засека или Дикротичен зарез е мало надолно отклонување во артерискиот пулс што ги одделува систолната и дијастолната фаза. Тоа е забележано во аортниот сигнал на притисок како последица на затворање на аортниот вентил.

За време на систолната фаза настанува зголемување на волуменот на крв во капиларното корито како резултат на брзиот проток на артериската крв, додека за време на дијастолната настанува намалување бидејќи крвта излегува од артериите. Ова обезбедува квалитативно мерење со слабеење на светлината низ ткивото како функција на времето, а тоа резултира во пулсирачки бран што ја претставува AC компонента на PPG-сигналот.

Карактеристиките како што се времето на пораст, амплитудата и формата можат да предвидат васкуларни промени во крвотокот. Исто така, можно е да се забележат две уникатни фази: растечки дел на пулсот, кој исто така се нарекува анакропски (систола) и паѓачки раб, наречен катакротски, што претставува дијастола. Дополнително, дикротичниот зарез е обично видлив во катакротската фаза. Овие карактеристики се објаснети во табела 3.1.

Табела 3.1: Карактеристикина PPG-сигналот

Table 3.1 Characteristic of PPG signal

PPG карактеристики	Објаснување
Систолни амплитуди (Systolic Amplitude)	Ги одразуваат AC варијациите на крвниот волумен околу местото на мерење
Област на пулсот (Pulse Area)	Вкупна површина под PPG кривата, во интервал помеѓу два систолни врва
Интервал врв до врв (Peak to Peak Interval)	Интервал помеѓу два систолни врва
Индекс на крутост на големи артерии (Large Artery Stiffness Index)	Временскиот интервал помеѓу систолниот и дијастолниот врв

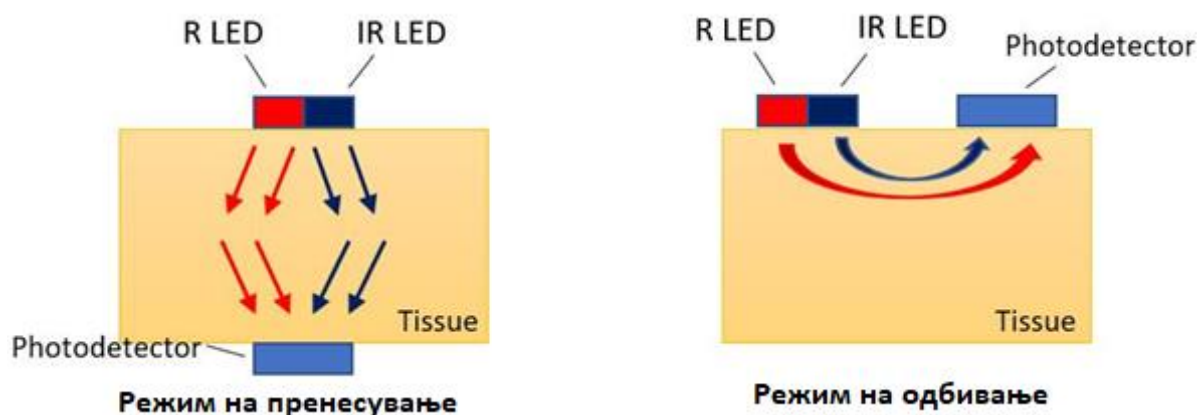
Другата компонента е обликувана од дишењето, активноста на симпатичкиот нервен систем и терморегулацијата. Додека AC компонентата има клиничка значајност заради нејзината синхронизација со срцевиот циклус. DC компонентата, исто така, обезбедува значајни информации, како што се состојби на хиперемија или хипоемија, температурни промени, венско волуменозни флукутации и други регулаторни механизми (Reisner et al., 2008).

Пулсниот оксиметар користи PPG сигнал за да извлече витални информации од субјектот со осветлување на дел од кожата и мерење на одбиената или пренесената светлина со помош на компонентата, фото-детектор. Ова, како што е спомнато и претходно, претставува неинвазивна техника за откривање на промените во волуменот на крвта во микроваскуларниот дел од ткивото (Allen, 2007). Главно има два вида пулсни оксиметри во зависност од режимот на работа, и тоа:

- Режим на пренесување.
- Режим на одбивање.

Како што е прикажано на слика 3.4, во режимот на пренесување, диодите што емитуваат светлина (LED) и компонентите на фотодетекторот што ја претвораат светлината во еквивалентни напони се монтирани на спротивната страна на мерната сонда. Повеќето клинички користени оксиметри се од видот кој работи на режимот на пренесување, каде што компонентата со фотодетектор ја собира светлината што се пренесува преку прстот, ушите или ножните прсти на

пациентот. Но, местата на кои се поставува, може да влијае и да пречи на извршувањето на секојдневните активности.



Слика 3.4: Видови на пулснооксиметри - режим на пренесување, режим на одбивање

Figure 3.4 Types of pulse oximeters – transmission mode, reflection mode

Во режимот на одбивање, фотодетекторите се на истата страна како и LED диодите. Ваквиот вид на пулсни оксиметри обично се монтираат на челото и се покажани како подобра алтернатива заради нивната можност да се постават на места на кои нивното носење не претставува пречка во извршување на секојдневните активности. Покрај тоа, PPG сигналите добиени на овој начин со користење на режимот за одбивање имаат подобар сооднос на сигнал кон шум (SNR - Signal to Noise ratio) во присуство на движечки артефакти. Затоа, и ние избравме да го користиме ваквиот вид на сензор, но наместо да биде поставен на челото тој е поставен да се носи околу зглобот на раката како часовник и на тој начин воопшто да не пречи во извршувањето на секојдневните активности (Johnston et al., 2004).

И покрај тоа што има широк опсег на LED диоди, што овозможуваат различна апсорпција на светлина од ткивото, зелената LED се смета за најчесто користена. Ова е само затоа што зелените светлосни бранови продираат во ткивото, создавајќи го најсилниот плетизмографски сигнал за мерење на светлосното одбивање и ја откриваат брзина на пулсот со повисок степен на прецизност (Nagre & Mendelson, 2005).

3.2. Движечки артефакти

Покрај светлосните бранови, најчесто можат да се јават и други странични влијанија што влијаат негативно на добивањето на сигналот од PPG сензорите. Овие странични ефекти што главно се добиваат како резултат на движење на субјектот се наречени движечки артефакти (Motion Artifacts - MA). На ваков начин се добиваат недоверливи информации и добивањето на точниот сигнал е намалено, а причина за тоа е движењето на субјектот или шумот кој се добива од самата околина. Поради ваквите движечки артефакти, PPG сензорот сè уште не е многу прифатен како сензор за апликации кај мобилни и здравствени апликации. Но, може да најде примена во повеќе клинички делови како: клиничко следење на физиолошки параметри (брзина на отчукување на срцето (HR), крвниот притисок, кислородна сатурација на крвта, респирација), васкуларна проценка - артериска усогласеност (Arterial Compliance), артериска болест и стареење, ендотелијална функција, вазо-спастична состојба, венска проценка) и автономна функција (терморегулација и вазомоторна функција, крвен притисок и варијабилност на отчукувањата на срцето, неврологија, ортостатска нетолеранција и други проценки на кардиоваскуларните промени).

Според (Chong et al., 2014), постојат три различни извори на движечки артефакти кои ја компромитираат читливоста на добиениот сигнал од PPG сензорот, а тоа се: еколошки, физиолошки и експериментални артефакти. Овие шумови можат да се припишат на мешањето на електромагнетните бранови околу телото; истовременото мерење и добивањето на други биофизиолошки сигнали или може да биде шум кој го создава самиот мерен инструмент.

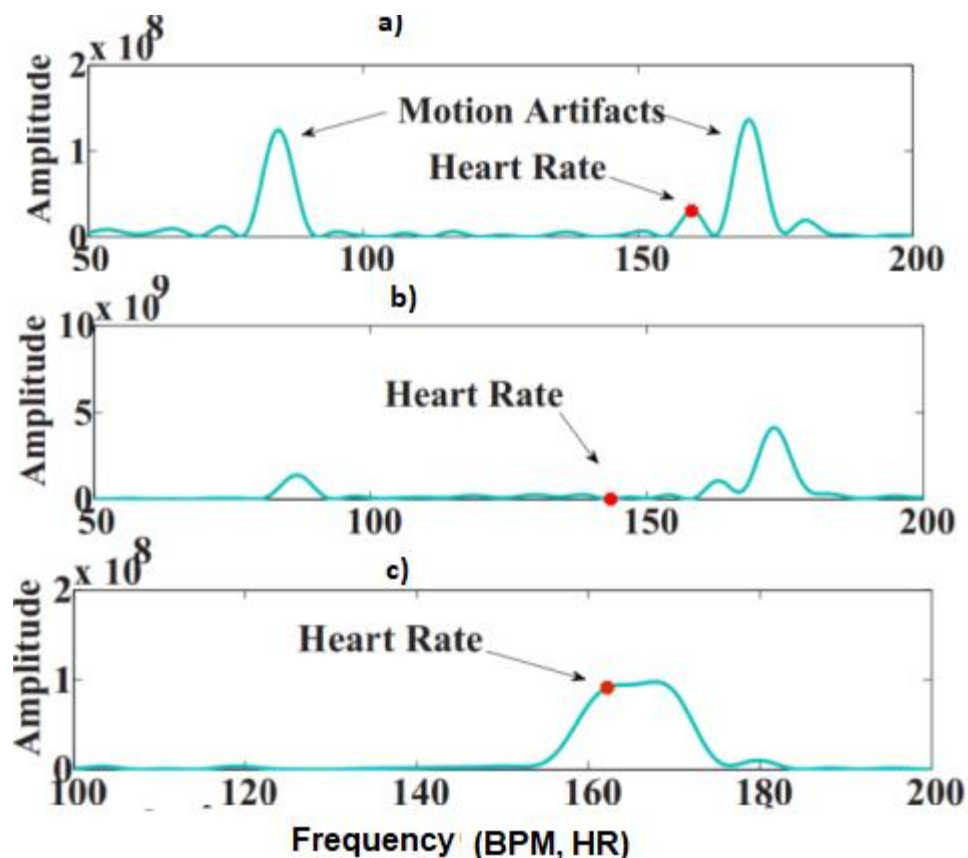
Други нарушувања во сигналот можат да бидат предизвикани од самата мерна локација, контактната сила при поставувањето на мерниот инструмент, шум добиен при механичко движење на телото или дишење, температура на околината, карактеристиките на кожата и сл.

Волните или неволните движењата на мускулите или дилатацијата на другите ткива може да направат промени во ткивото, а со тоа и на сигналот кој се добива. Светлината што се враќа од сензорот, може да се модифицира со ваквите движења, на тој начин се создава шум, а со тоа се нарушува и точниот сигнал.

Самата различна анатомија на различните субјекти, димензијата и големината на органите и количината на течности задржани од страна на ткивата, резултира со варијации во ширењето на светлината во внатрешните органи.

Друг фактор што го менува сигналот е релативното поместување на сензорот со кожата. Ова е предизвикано пред сè затоа што системот е чувствителен на движење, а импедансата на електрода-кожа се менува и при движење. За PPG сензорите кои се носат на рачен зглоб, за време на движењето на рацете, јазот помеѓу кожата на зглобот и сензорот се менува и како резултат на тоа, измерениот интензитет на PPG-сигналот варира со него. Ваквата варијација е често поврзана со фреквенцијата во која се движи раката.

Како што се гледа на слика 3.5 а), PPG-сигналот е корумпиран од МА заради движење и физичко поместување помеѓу сензорот и кожата, што настанува при физичко движење или трчање. Во овој случај вистинскиот врв на HR и врвот на движечките артефакти се раздалечени еден од друг. Може да се случи сензорот толку многу да се раздели од кожата, па оригиналниот врв што одговара на HR може да биде целосно отсутен од спектарот. Слика 3.5 б) го покажува овој случај. Понекогаш, фреквенцијата на движење и фреквенцијата на HR може да се приближат до тоа што HR станува незабележлив во однос на шумот, за дадена резолуција на периодограмот, како што е прикажано на слика 3.6 с).



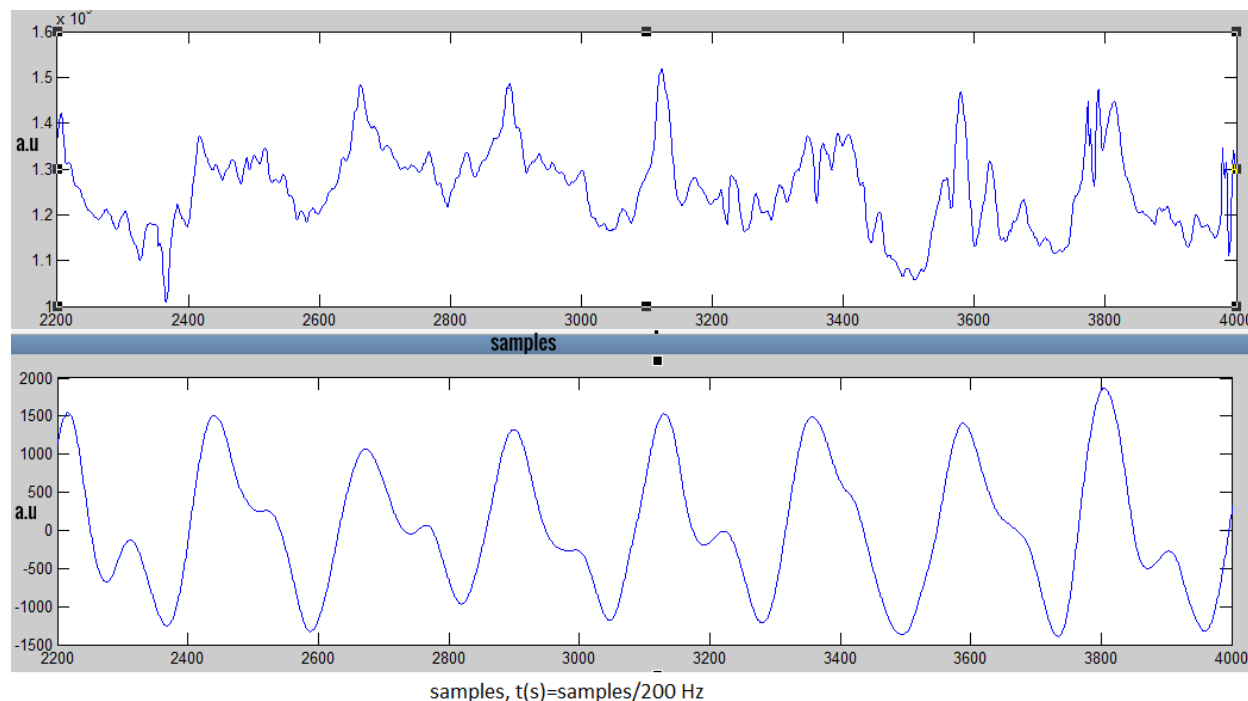
Слика 3.5: Периодограм на PPG-сигналот а) Случај кога највисоките вредности на HR и на шумот се наоѓаат на поголемо растојание б) Случај кога врвот на HR е целосно отсутен в) Случај кога врвот на HR не може да се определи

Figure 3.5 Periodogram of PPG signal. A) Case where the highest values of HR and noise are at a greater distance. b) Case where the top of the HR is completely absent. c) Case where the peak of the HR cannot be determined

3.3. Дефиниција на проблемот

Постојат бројни производи кои користат PPG сигнали за следење на повеќе биофизиолошки параметри, но тие се далеку од прецизни. Главниот проблем со ваквите уреди е тоа што тие се многу осетливи на движечките артефакти – МА, кои се спомнати погоре. На слика 3.6 е прикажан пример од PPG сигнали, со и без движечки артефакти, соодветно (на апцисата е претставен бројот на примероци, бидејќи фреквенцијата на испитуваните сигнали е 200Hz, односно 200 примероци

во секунда, тоа значи дека на апцисата се преставени 9 секунди од сигналот – $t(s)=samples/200Hz$).



Слика 3.6: PPG сигнал со и без движечки артефакти

Figure 3.6 PPG with and without moving artifacts

Движечките артефакти се јавуваат како резултат на движењата на рацете, кога станува збор за уреди кои PPG-сигналот го добиваат со носење на сензорите на зглобот на раката. Ваквите движења на раката може да влијаат на вредностите од PPG сигналите. Затоа нашата главна цел е да добиеме алгоритам кој го отстранува шумот, посебно шумот добиен при движење на PPG-сигналот. На тој начин од прочистениот PPG со едноставна пресметка може, со голема прецизност, да се естимира брзината на отчукување на срцето, а со тоа и да се определат други биофизиолошки параметри, на пр. крвниот притисок.

Денес постојат многу уреди кои се користат за набљудување на човековото здравје. Една од особеностите и можностите која ја имаат најголемиот дел од нив е следење на срцевата работа, односно брзината на отчукувања, што најчесто се прави со користење на PPG сигнали, при што мерењето, како што веќе објаснивме, наједноставно е да се врши на зглобовите на рацете. проблемот се состои во тоа што мерењата на срцевата работа може да бидат „искривени“

поради движењата кои настануваат кога лицето прави некои физички движења. Ова го прави набљудувањето на срцевата работа многу покомплексно, но последователно на тоа и помалку точно. Ова е причината поради што се јавува потреба од алгоритам кој би ги надминал ваквите проблеми.

Обиди за решавање на проблемот, за да се извлечат биофизиолошки параметри, на пр., HR, од PPG сигнали, како и алгоритми кои ги отстрануваат шумовите создадени кај PPG сигналите, има повеќе. Оваа проблематика особено е актуелна во последните неколку години (M. Wang et al., 2019; Yifan Zhang et al., 2019). Добро познатата техника, TROIKA (Z. Zhang et al., 2015) се состои од три главни делови. Тоа се декомпозиција на сигналот за отстранување на шум, ретка реконструкција на сигналот за проценка на спектарот со висока резолуција и следење на спектралниот врв со механизми за верификација. Во овој метод, грешката е драстично намалена во однос на претходните истражувања, дури и во присуство на движење со голема брзина.

Друг неодамнешен начин за следење на срцевиот ритам (Khan et al., 2016) покажа подобар резултат од TROIKA при движење со голема брзина. Во оваа нова техника, почетната фаза на одмор или претходно пресметаната HR вредност не е потребна. Овде, апсолутна состојба на критериумот е воведена врз основа на декомпозицијата на емпирискиот режим (ensemble EMD - EEMD). Поради користење на два канални PPG сигнали, алгоритамот е стабилен. Покрај тоа, RLS филтрирањето и екстракција на временскиот домен ги намалуваат грешките и ја зголемуваат стабилноста на овој алгоритам.

Исто така, од не толку одамна предложен е и друг алгоритам (Jarchi & Casson, 2017) за да се процени моменталната фреквенција на Хилбертовата трансформација на сигналот PPG за време на интензивно движење. Иако грешката проценета од овој алгоритам не е помала од оние прикажани во (Khan et al., 2016) и (Z. Zhang et al., 2015), сепак инстантното HR може да обезбедат повеќе информации за човечкото тело. Инстантниот ритам на срцето може да биде директен показател на активноста на вагалниот нерв и симпатичен нервен систем.

Методот предложен (во Z. Zhang, 2015) проценува спектар на PPG сигнали заедно со користење на повеќе мерни вектори моделирани со ретка реконструкција на сигналот (sparse signal recovery). Ова лесно може да ги идентификува и отстрани спектралните врвови на движевките артефактите, како резултат на општото ограничување на реткост кај спектралните коефициенти. Прикажано е дека работи добро при ниски рати на семплирање.

Друг алгоритам е предложен во (Fujita et al., 2018), каде се прави естимација на HR базирана на парцијален филтер (Particle filter-based heart rate estimation -PARHELIA) со користење на PPG сигнали за време на физичка вежба што го намалува времето за проценка на HR, а со тоа покажува подобри резултат во однос на (Z. Zhang, 2015). Ваквиот филтер го естимира HR во мал опсег на податоци каде што временскиот период е краток. Бидејќи HR варијациите се мали во мал опсег на податоци, тој покажува подобри резултат (Fujita et al., 2018).

Техниката предложена (во Tanweer et al., 2017) за мерење на HR се смета дека е погодна за прекумерна појава на движевчки артефакти во сигналот PPG. Овде се користат податоци за забрзување по 3-те оски, а за намалување на движевките артефакти се користи повеќестепениот X-LMS филтер. Филтерот со повеќе фази X-LMS означува дека излезот на адаптивниот филтер е вратен на влезот неколку пати за да се намали грешката. За следење на врвовите на HR, во овој алгоритам се користи методот Slope Sum Method (SSM).

Алгоритамот предложен (во Baca et al., 2015) служи за проценка на HR преку отстранување на движевките артефакти (MAR) и прилагодено следење (adaptive tracking -AT). Овде за декомпозиција на сигналот се користи декомпозиција со единичната вредност (Singular value decomposition -SVD) за тој да се раздели во два потпростора. Овде се користи FFT за наоѓање на врвови (FFT peak finder) за да се добијат доминантните фреквенции. Овој алгоритам е поедноставен во споредба со другите, но проценетата грешка е голема.

Друг алгоритам е предложен во (Krishnan et al., 2010) за да се отстрани MA со анализа на податоците од две фази. Во првата фаза, податоците се анализираат со статистичка анализа од повисок ред, со која се утврдува

присуството на МА, а во втората фаза се намалуваат постоечките движечки артефакти.

Алгоритмот развиен (во Lopez-Silva et al., 2012) е заснован на линеарно филтрирање, работи во фреквентен домен и користи евристичка анализа. Во овој алгоритам, за првпат се развива евристичкиот пристап со набљудување на PPG-сигналот. Овој алгоритам најпрво идентификува три фреквенции од PPG-сигналите, каде што две МА можат да се идентификуваат со едноставна логика, а преостанатата фреквенција одговара на HR врвот. Овој алгоритам не работи добро при интензивно вежбање.

За првпат ефектите на ткивата се вклучени при пресметката HR во (Yousefi et al., 2014), со употреба на пулсниоксиметри. Овде за отстранување на МА и намалување на шум од друга природа се користи адаптивен отстранувач на шум во две фази.

Главното, сите досегашни начини на кои се работи можат да се поделат во три главни категории (Yifan Zhang et al., 2019):

- Методи со користење на единичен PPG сигнал.
- Методи со користење на повеќе каналски PPG сигнали.
- Методи со користење на PPG сигнал заедно со податоци за движењето, најчесто податоци за забрзување по 3-те оски.

Алгоритмите кои во основа работат со единичен PPG сигнал, генерално, користат методи на декомпозиција, временска или фреквентна, за да ја одделат МА компонентите од корумпираниот сигнал. На пример, (Z. Zhang et al., 2015) користи спектрална анализа за декомпозиција на PPG-сигналот на повеќе компоненти. Покрај овој метод на декомпозиција на сигналот други користат декомпозиција на емпириски режим (empirical mode decomposition - EMD), wavelet декомпозиција, единичен мерен вектор (single measurement vector - SVM) и повеќекратни мерни вектори (multiple measurement vector - MMV) (C. M. Lee & Zhang, 2003). Генерално, овие алгоритми работат релативно добро со непериодични движечки артефакти, но кога движењата се повторуваат и посебно ако станат интензивни, точноста во пресметките вообичаено се намалува, бидејќи во таков случај станува многу тешко разликувањето на PPG-сигналот од движечките артефакти.

Некои од обидите за отстранување на ваквиот шум, кои користат повеќе каналски PPG сигнали, како што е на пример методот за независна компонентна анализа (ICA - Independent Component Analysis), ги одделува движечките артефакти во временски домен, користејќи претпоставка дека МА и чистиот PPG сигнал се статистички независни. Но, ваквите сигнали не секогаш, се статистички независни, бидејќи движењата можат да влијаат врз артерискиот проток, а тоа да влијае на PPG. Бидејќи во такви случаи, ваквите алгоритми не се ефикасни, предложен е друг метод (во Yousefi et al., 2014) (Fallet & Vesin, 2017) (Chowdhury et al., 2018) (Lei Wang et al., 2007). Кај (Yousefi et al., 2014) се користат инфрацрвени и црвени канали на оксиметар поставен на прстот од раката, во (Chowdhury et al., 2018) и (Fallet & Vesin, 2017) се користат два зелени канали поставени на зглобот на раката, во (Lei Wang et al., 2007) се користат два инфрацрвени канали поставени на увото, а во (H. Lee et al., 2018) се користат дури 9 канали за читање на PPG. Во некои од истражувањата се користат дополнителни податоци за забрзување по 3-те оски (Z. Zhang et al., 2015)(Z. Zhang, 2015) (Təuṭan et al., 2015).

Неодамна е истражена и употребата на жirosкоп како показател за движење. Според овие истражувања жirosкопот дозволува одвојување на забрзувањето заради движењето на сензорот или поради гравитацијата и може да се користи како референца за движење при отстранување на МА (Casson et al., 2016) (Casson et al., 2016)(H. Lee et al., 2019). Алгоритмите во оваа категорија работат и со временски и фреквентен домен. Кога се користи временскиот домен, најчесто се користат прилагодливи методи за филтрирање како што се Least Mean Square (LMS) и Recursive Least Squares (RLS). А кога се работи во фреквентен домен, одземањето на спектарот (Spectrum Subtraction - SS) претставува широко прифатен пристап. Овие методи многу зависат од големата корелација помеѓу МА во сигналот PPG и податоците за движење (Ye et al., 2017)(Ram et al., 2012).

3.4. Нашиот алгоритам

Ние како темел ги земаме принципите кои ги поставува околината TROIKA, користиме делови од тој алгоритам, го прилагодуваме за да го применуваме на нашето множество од податоци. За разлика од поголем број на алгоритми кои

погоре ги спомнавме, кои користат готово множество на податоци за тренирање и тестирање, ние користиме сопствено множество на податоци. Нашето множество на податоци се состои од мерења на PPG сигнали врз лица кои се згрижени во старечки дом. Сензорите за собирање на PPG-сигналот се поставени на зглобот од раката. Уредот работи со фреквенција од 200 Hz, што значи дека на секои 0.005 секунди се мери нова вредност за PPG. Бидејќи податоците се собираат од дом за згрижување на стари лица, кај нив нема интензивни вежби и движења, сепак движечките артефакти многу лесно можат да направат проблем во сигналот.

TROIKA алгоритмот се состои од повеќе делови: првична декомпозиција на сигналот (Signal Decomposition), реконструкција на сигналот (Sparse Signal Reconstruction - SSR) и следење на спектралниот врв (Spectral Peak Tracking - SPT). Овој алгоритам работи со единичен PPG сигнал и користи податоци за забрзување. Ова одговара на нашите потреби бидејќи и ние во собраните податоци имаме единичен PPG сигнал и податоци за забрзување. Пред да се започне со фазата на декомпозиција, PPG-сигналот минува низ филтер на опсег. Ваквиот филтер ги отстранува фреквенциите кои не се поврзани со отчукување на човеково срце.

Фазата на декомпозиција на сигнал се користи за отстранување на шумот и движечките артефакти од PPG-сигналот. Ние во ова поглавје се задржуваме токму на овој дел. Ова се прави со декомпозиција на единичниот сигнал во повеќе компоненти. Ваквите компоненти потоа се анализираат и оние кои се поврзани со шум и одредена интерференција може да бидат отстранети. Потоа, сигналот се реконструира без тие компоненти. Пред започнување на фазата SSR, сигналот привремено се диференцира заради отстранување на MA компонентите кои кореспондираат со апериодични движења. SSR фазата обезбедува проценка на сигналот со висока резолуција, што го прави поедноставно следењето на врвовите. SPT фазата го анализира сигналот креиран од SSR фазата и ги селектира спектралните врвови кои одговараат со срцево отчукување. Притоа се користи хармоничната релација на фреквенцијата на срцевата работа и фактот

што во две последователни временски рамки не може да постои голема промена во срцевата работа.

3.4.1. Алгоритми за декомпозиција на сигналот

Како што рековме, во ова поглавје ние се фокусираме на добивање на прочистен PPG сигнал, односно сигнал без шум, од кој понатаму лесно можат да се извечат потребните карактеристики, затоа во овој чекор се задржуваме до делот за декомпозиција на сигналот. Некои од алгоритми за декомпозиција на сигналот, кои и погоре се спомнати се Singular Spectrum Analysis (SSA), Single-Channel Independent Component Analysis (SCICA) и Empirical Mode Decomposition (EMD).

SSA е моќна техника за анализа на временски интервали (Time series analysis), која често и успешно се користи во ваквата проблематика. Може да се користи со многу апликации и работи со произволни статистички процеси. Ова вклучува линеарни и нелинеарни, стационарни и нестационарни сигнали итн. SSA врши декомпозиција на временските интервали во осцилаторни компоненти и шум. Недостатокот на SSA алгоритмот се состои во тоа што тој не врши автоматско отстранување на MA. Ова е поради тоа што нема доволно информација за да се утврди кои компоненти го содржат сигналот од срцевото отчукување. Сепак, постои решение за ова, а тоа се акцелерометарските податоци, односно податоците за забрзување, да се користат за препознавање на MA компонентите во SSA декомпозицијата, а потоа тие да бидат исклучени во процесот на реконструкција.

SCICA алгоритмот може да се користи за издвојување на независни примероци од единични временски интервали. Тој го користи ICA алгоритмот за поделба на единичен канал во повеќе компоненти. ICA е техника за поделба на линеарно помешани извори. Недостаток на SCICA алгоритмот е тоа што за процесите од независните извори потребно е да имаат раздвоена спектрална поддршка за разделување на спектарот. Тоа не е случај кај ЕКГ и PPG сигналите, затоа потребни се повеќе информации за сигналот и на тој начин потешко е да се направи поделба на истиот.

EMD алгоритмот е емпириски алгоритам, кој е повеќе интуитивен. Тој функционира добро на нестационарни и непериодични сигнали, какви што се и

нашите. EMD прави декомпозиција на PPG-сигналот и сигналите за забрзување во внатрешна режим функција (intrinsic mode function - IMF) Недостаток на EMD алгоритмот е тоа што не може да ги идентификува оние IMF чии фреквенции се многу блиску едни до други. Ова може да претставува проблем при одредување на пулсот со употреба на PPG монитор околу рачниот зглоб. А нашиот е токму таков случај. Доколку фреквенцијата на МА е близу до фреквенцијата на срцевото отчукување, тоа ќе влијае на крајниот резултат. Но, според (R. & Prabhu, 2015) овој проблем може да биде решен со употреба на малку изменета верзија на EMD алгоритмот, наречен EEMD (Ensemble EMD). Оваа верзија додава бел шум на оригиналниот сигнал секогаш пред да помине низ EMD. Разгледувајќи ги недостатоците на секој од алгоритмите, SSA алгоритмот претставува најветувачка техника, прилагодлива за нашите податоци. Недостатокот кој го спомнавме може лесно да се реши, бидејќи податочното множество ги вклучува и податоците за забрзување.

3.4.2. Сингуларна спектрална анализа (Singular Spectrum Analysis)

Еден од алгоритмите за декомпозиција на сигналот што ние одлучивме да го користиме за нашиот случај е SSA (Singular Spectrum Analysis) алгоритмот. Овде ќе објасниме како работи алгоритмот и неговата имплементација во MATLAB. Се користат следните ознаки:

- Со задебелен (Bold) фонд и голема буква означуваме матрица (пример: X).
- Со стандардни големи букви означуваме вектор (пример: X).
- Со нормални мали букви означуваме скалари (пример: x).
- Горен индекс Т се користи за транспонирање (пример: X^T).
- Индексите i, j се користат за означување на елементи во матрицата, редица i и колона j (пример: $x_{i,j}$)

Имплементирањето на SSA алгоритмот е прилично јасно: потребно е следење на чекорите на имплементација. Основниот SSA алгоритам се состои од две комплементарни фази: декомпозиција и реконструкција. Чекорите во фазата на декомпозиција се дадени во (Hassani, 2007):

1. Пресметување на X матрицата на траекторијата - L . Ова ги трансформира еднодимензионалните временски низи $Y_N = (y_1, \dots, y_N)$ во мултидимензионални низи X_1, \dots, X_K со вектори $X_i = (y_i, \dots, y_{i+L})^T$. L е должината на прозорецот и е избран така што $2 \leq L \leq N$. K е избрано така што $K = N - L + 1$. Резултатот е матрицата на траекторијата $X = [X_1, \dots, X_K]$:

$$X = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_N \end{pmatrix} \quad (3.1)$$

2. Се пресметува матрицата $S = XX^T$. Ова е позитивно дефинирана и симетрична матрица на која може да и се направи декомпозиција со користење на декомпозиција со сопствени вредности. Исто така, сопствените вредности на оваа матрица се реални и позитивни.
3. Декомпозиција со единична вредност вредносните (Singular value decomposition SVD) на матрицата S . Прво се пресметуваат сопствените вредности и сопствените вектори на матрицата S . Ако $\lambda_1, \dots, \lambda_d$ се сопствени вредности на S во ред на намалување и U_1, \dots, U_d се соодветните сопствени вектори од сопствените вредности, тогаш $V_i = X^T U_i / \sqrt{\lambda_i}$ ($i = 1, \dots, d$). Овде $d = \max\{i, \text{така што } \lambda_i > 0\}$. Матрицата на траекторијата сега може да се запише како:

$$X = X_1 + \dots X_d, \quad (3.2)$$

каде $X_i = \sqrt{\lambda_i} U_i V_i^T$. Колекцијата $(\sqrt{\lambda_i}, U_i, V_i^T)$ се нарекува сопствена тројка на SVD.

Стапките на фазата за реконструкција се:

1. Групирање: Елементарните матрици X_i се разделуваат во m разделни подмножества I_1, \dots, I_m . Нека $I = i_1, \dots, i_p$, тогаш матрицата X_I одговара на

групата I и е дефинирана како $X = X_I(1) + \dots X_I(p)$. Матрицата на траекторијата X може да се запише како:

$$X = X_I(1) + \dots X_I(m). \quad (3.3)$$

2. Дијагонален просек: Секоја матрица X_I се трансформира во нова временска серија со должина N . Ова се компонентите на која е поделена оригиналната временска серија. Нека X биде $X \times L$ матрица со елементи $x_{i,j}$, $1 \leq i \leq L, 1 \leq j \leq K$. Тогаш $L^* = \min(L, K), K^* = \max(L, K)$ и $N = L + K - 1$. Нека $x_{ij}^* = x_{ij}$ ако $L < K$ тогаш $x_{ij}^* = x_{ji}$. Дијагоналниот просек трансформира матрицата X во серии g_0, \dots, g_{N-1} преку следната формула:

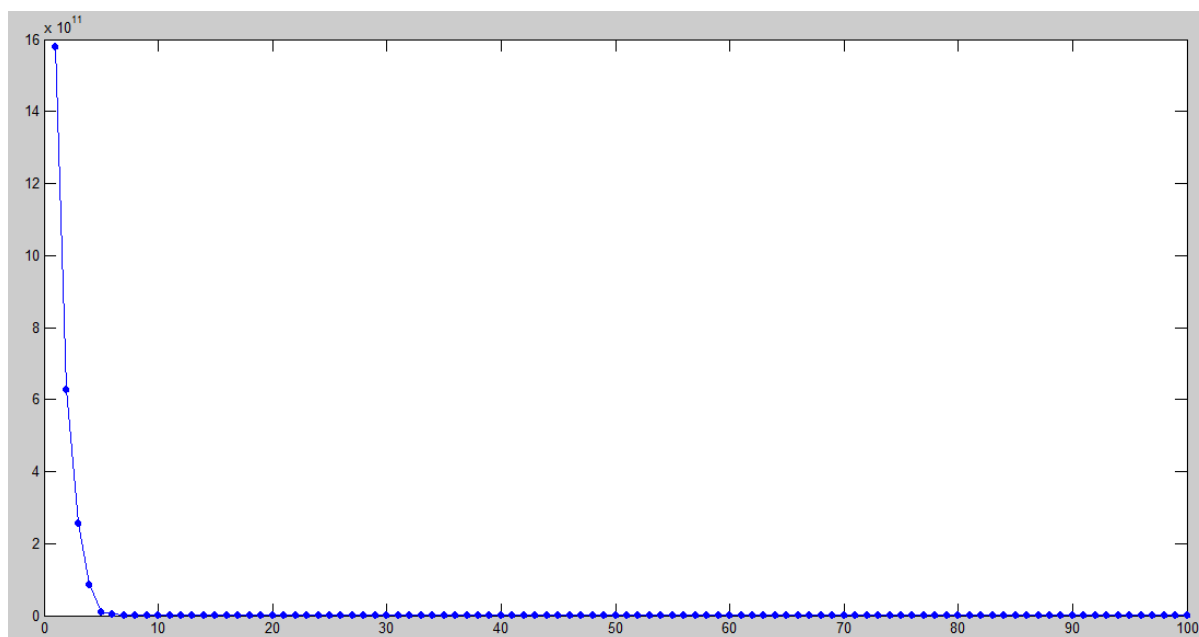
$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} x_{m,k-m+2}^* & 0 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} x_{m,k-m+2}^* & L^* - 1 \leq k < K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} x_{m,k-m+2}^* & K^* \leq k < N. \end{cases} \quad (3.4)$$

3.4.3. Оптимизирање и прилагодување на параметрите

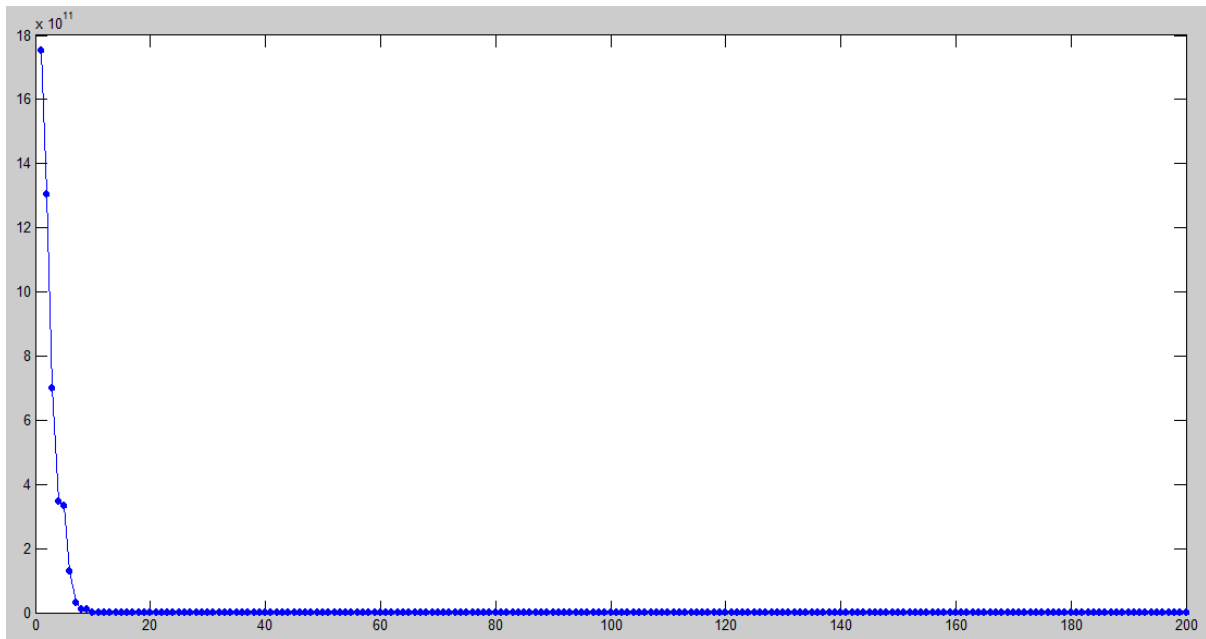
Алгоритмот SSA има два параметри кои може да се користат за оптимизација на сигналот врз кој се прави декомпозиција. Тоа се должината на прозорецот L и изборот на групирање. Генерално за должината на прозорецот L добро е да се избере вредност што е помала од половината од должината на сигналот што се декомпозира, така што да може да се избалансира грешката и се создава можност да се оперира со пониски фреквенции. Исто така, ако имаме некоја форма на периодичност во сигналот, L е избран за да биде делител на периодот од сигналот. Групирањето е направено на начин во кој сопствените вредности кои имаат амплитуда со иста големина да се групираат заедно.

Кај матрицата на траекторијата X во равенството (3.1) може да се види дека должината на прозорецот е висината на траекторијата на матрицата.

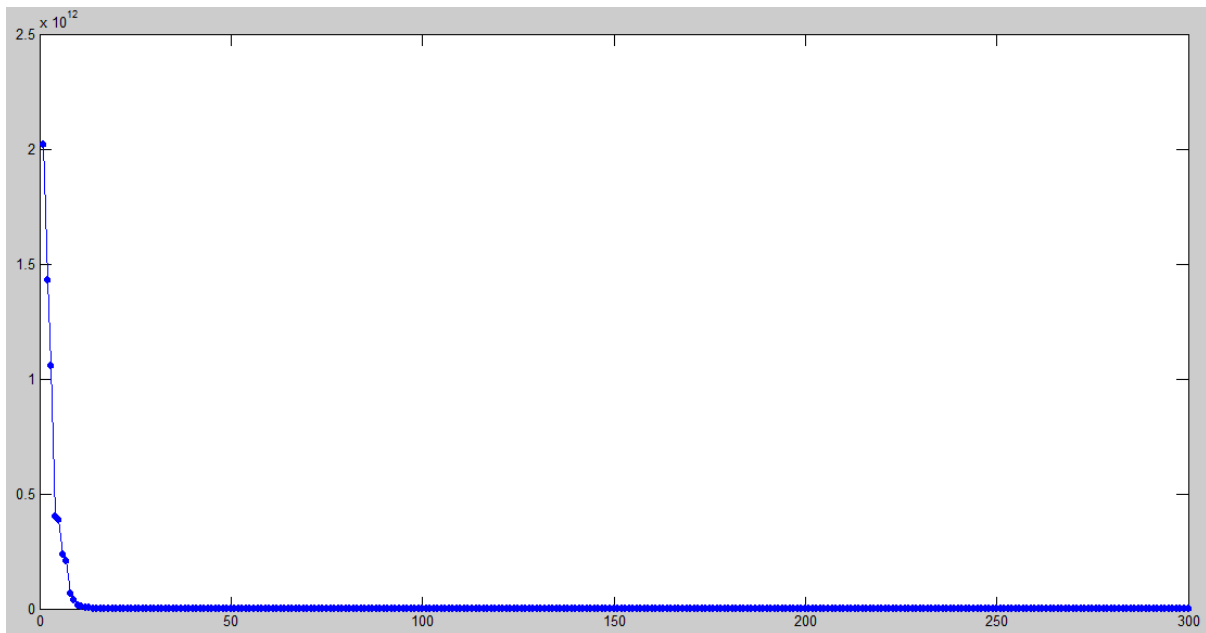
Количеството информација која се содржи во векторите X_i зависи од должината на прозорецот. SVD се пресметува за матрицата XX^T со димензии $L \times L$ така што должината на прозорецот одговара на бројот на сопствените вредности на матрицата. За да се избере размерот на должината на прозорецот, сопствените вредности мора да бидат разгледани исто така. На слика 3.7 се прикажани четири дијаграми на сопствени вредности на матрицата XX^T со различни должини на прозорецот.



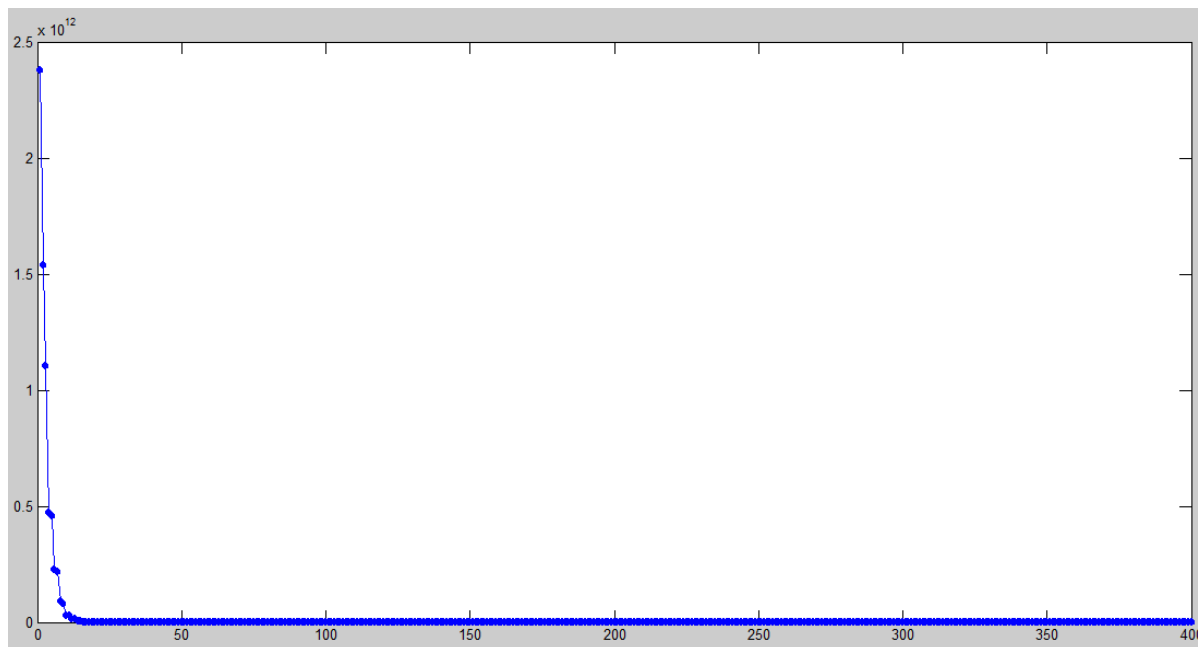
а) должина на прозорецот 100
a) window length 100



b) должина на прозорецот 200
b) window length 200



c) должина на прозорецот 300
c) window length 300



d) должина на прозорецот 400

d) window length

Слика 3.7: График на сопствените вредности

Figure 3.7 Graph of eigen values

При разгледување на четирите графикони на слика 3.7 може да се види дека размерот на должината на L навистина кореспондира со количината на сопствените вредности на матрицата XX^T . Размерот на сопствените вредности e е во релација со сумата на сите сопствени вредности од y -оската и се пресметува на следниот начин:

$$e_i(\text{relative}) = \frac{e_i}{\sum_{m=1}^L e_m} \times 100. \quad (3.5)$$

Ова е направено за да се види колку информација имаат сопствените вредности за оригиналниот сигнал. При споредување на слика 3.7 а) со слика 3.7 б), може да се забележи дека на слика 3.7 а) има помалку сопствени вредности кои имаат релативно поголеми вредности (поголем од 5%). Ова значи дека некои информации се отстранети кога L е избрано да биде премногу мало. Споредувајќи ја слика 3.7 б) со слика 3.7 с), исто така, може да се забележи дека сè уште има мал број на сопствени вредности слика 3.7 б) кои носат повеќе битни

информации споредени со слика 3.7 c). Со споредба на слика 3.7 c) и слика 3.7 d) се покажува дека количината на големи сопствени вредности се исти, но дека има повеќе сопствени вредности кои се нула или приближни до нула. Ова води до заклучок дека голем размер на прозорецот не содржи многу информации при реконструкција на сигналот. Земајќи го предвид сето ова, ефектите од реконструкција на сигналот ќе бидат истражени во интервалот за $200 \leq L \leq 300$. Во фазата на групирање матриците кои припаѓаат на сопствените тројки каде сопствените вредности се близу еден до друг по големина се групираат заедно. Ова е направено поради тоа што матриците основно се еднакви помеѓу себе. Сопствените тројки кои имаат вредности кои се близу до нула не содржат значајни информации и тие можат да бидат исклучени од реконструкцијата. Ова, исто така, го намалува обемот на процесирање, поради тоа што кореспондирачките матрици во фазата на групирање повеќе не се пресметуваат.

Според емпириски анализа, врз нашите податоци, добиваме дека најдобри резултати гледано за целото множество на податоци, се добиваат за должина на прозорецот $L=290$ и бројот на сопствени вредности да е 11. Ова, исто така, дава најдобри податоци во односот обем на процесирање и точноста на реконструкцијата.

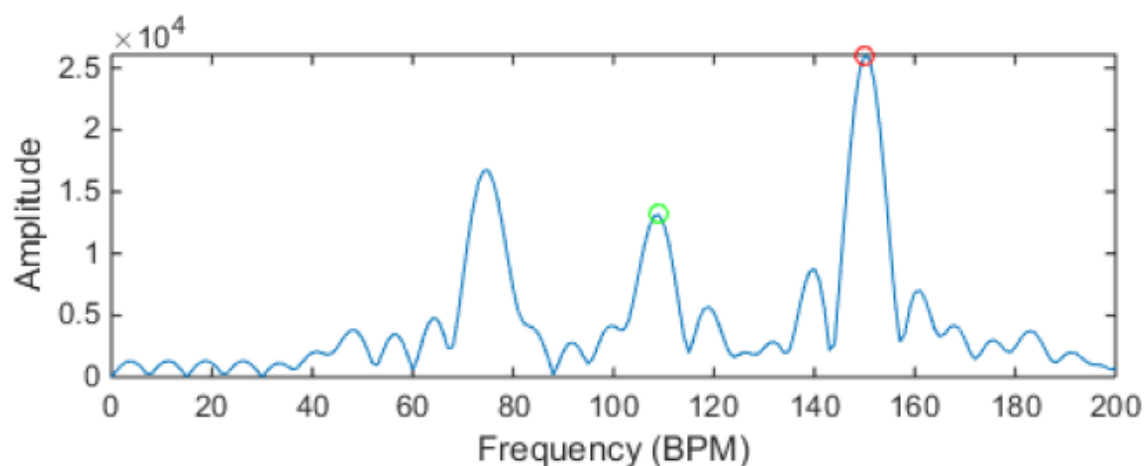
3.4.4. Отстранување на движечки артефакти

Како што веќе споменавме, најголемата причина за дисторзија на PPG-сигналот се движечките артефакти. Со SSA алгоритмот и декомпозиција на PPG-сигналот, тој може да биде поделен во повеќе компоненти, а една или повеќе компоненти содржат информација за срцевото отчукување. Сепак, идентификувањето на ваквите компоненти не е секогаш едноставно. Всушност, не е секогаш случај, првите компоненти да ја содржат информацијата за пулсот/срцевото отчукување. Која било од компонентите може да ја содржи таа информација. Токму поради ова, постои потреба од метода за селекција на вистинските компоненти. Потребно да бидат земени во превид бројни ситуации, но и да бидат поставени одредени критериуми за да бидат идентификувани бараните компоненти.

Начините на кои МА се појавува во PPG сигналите не се секогаш конзистентни. Можат да се јават и вакви случаи:

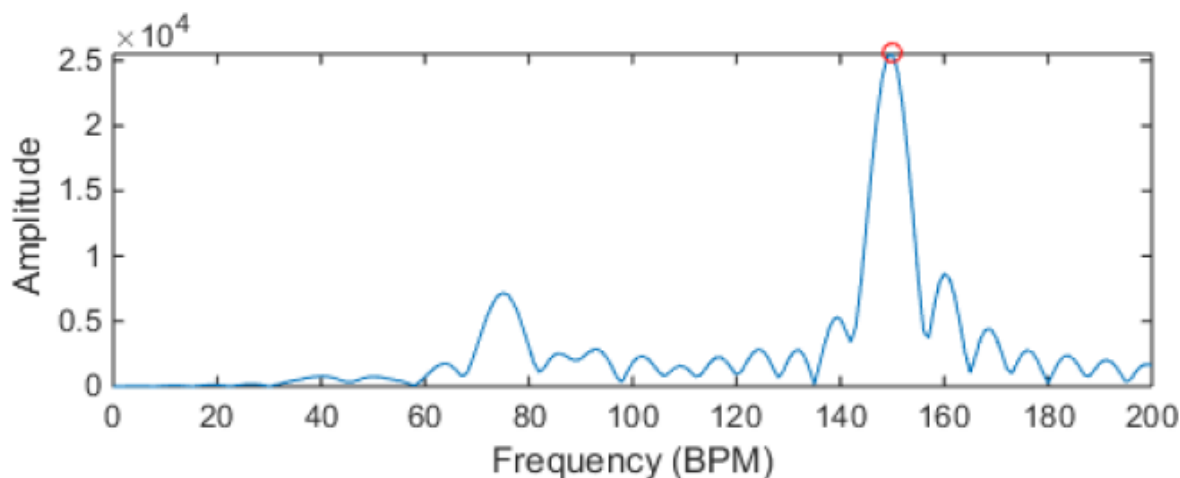
- Фреквенциите на МА врвовите се далеку од оние на срцевото отчукување (HR).
- Фреквенциите на МА врвовите се совпаѓаат со тие на PPG.
- Постојат нагласени МА врвови, а заради тоа може да отсуствува информација за врвовите во PPG-сигналот.
- Не постои јасен врв на PPG.
- Фреквенцијата на МА врвот е многу близу до онаа на PPG.

На слика 3.8 се дадени различни графикони за приказ на наведените ситуации.



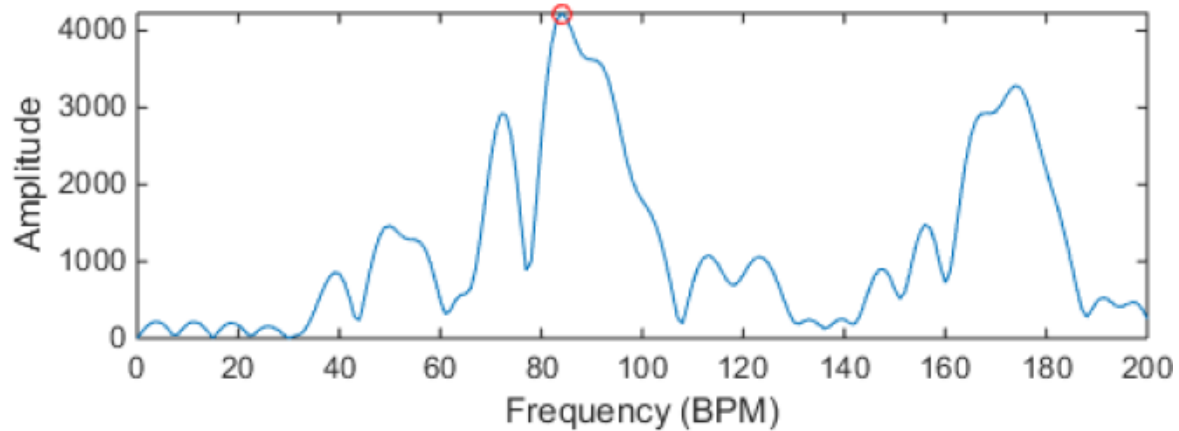
a) фреквенциите на МА врвовите се далеку од оние на HR

a) MA peaks frequencies are far from those of the HR



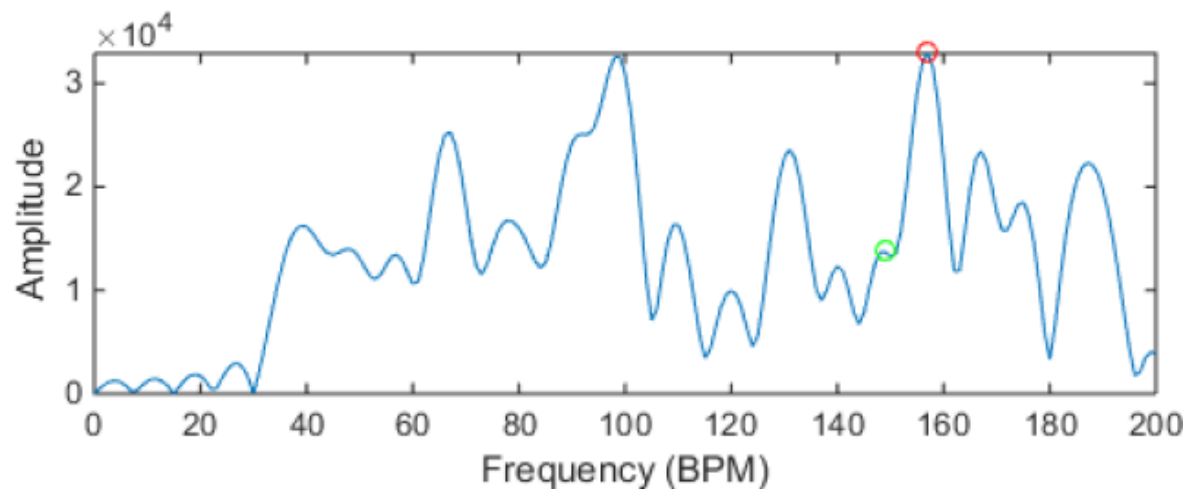
b) фреквенциите на МА врвовите се совпаѓаат со тие на PPG

b) MA peaks frequencies match with those of the PPG



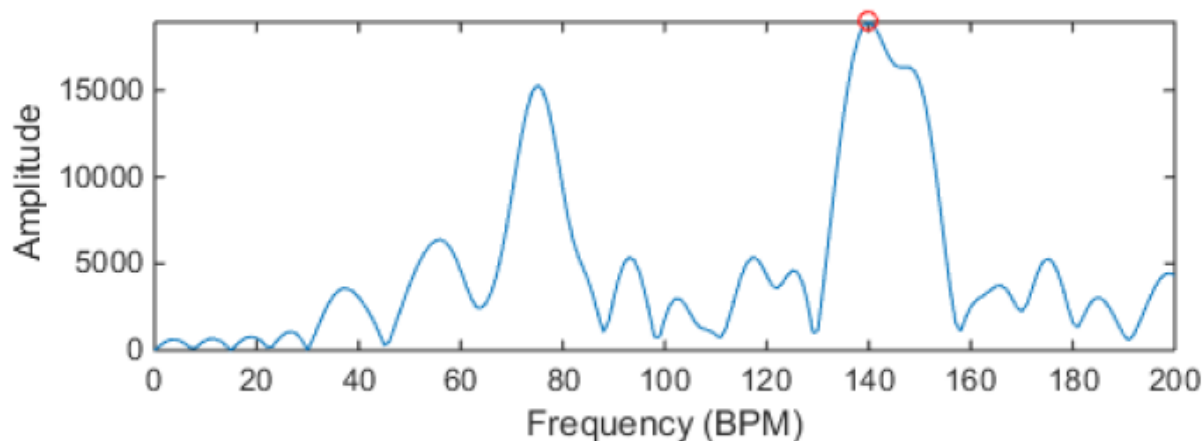
с) постојат нагласени МА врвови, а заради тоа може да отсуствува информација за врвовите во PPG-сигналот

c) there are emphasized MA peaks, and therefore may to removes information on peaks in the PPG signal



d) не постои јасен врв на PPG

d) there is no clear PPG peak



е) фреквенцијата на МА врвот е многу близу до онаа на PPG

е) MA peak frequency is very close to that of the PPG

Слика 3.8: Начините на кои движечките артефакти можат да се појават во PPG
сигналите

Figure 3.8 The ways when moving artifacts can appear in PPG signals

Црвената и зелената точка во графиконите ги означуваат максималната вредност и вистинската вредност за HR. Доколку нема зелена точка на графиконот, тоа значи дека највисоката вредност се совпаѓа со HR. PPG сигналите и податоците за забрзување се конвертирани во фреквентен домен преку брза фуриева трансформација (FFT). Ова е направено за да се направи поточна разлика меѓу врвовите. Друго што треба да се потенцира овде е тоа што врвовите во податоците за забрзување се исто така прикажани во PPG-сигналот. За справување со МА, ние се задржуваме на делот за отстранување на компоненти кои се вклучуваат во самиот дел за декомпозиција (Removing of Components-RoC).

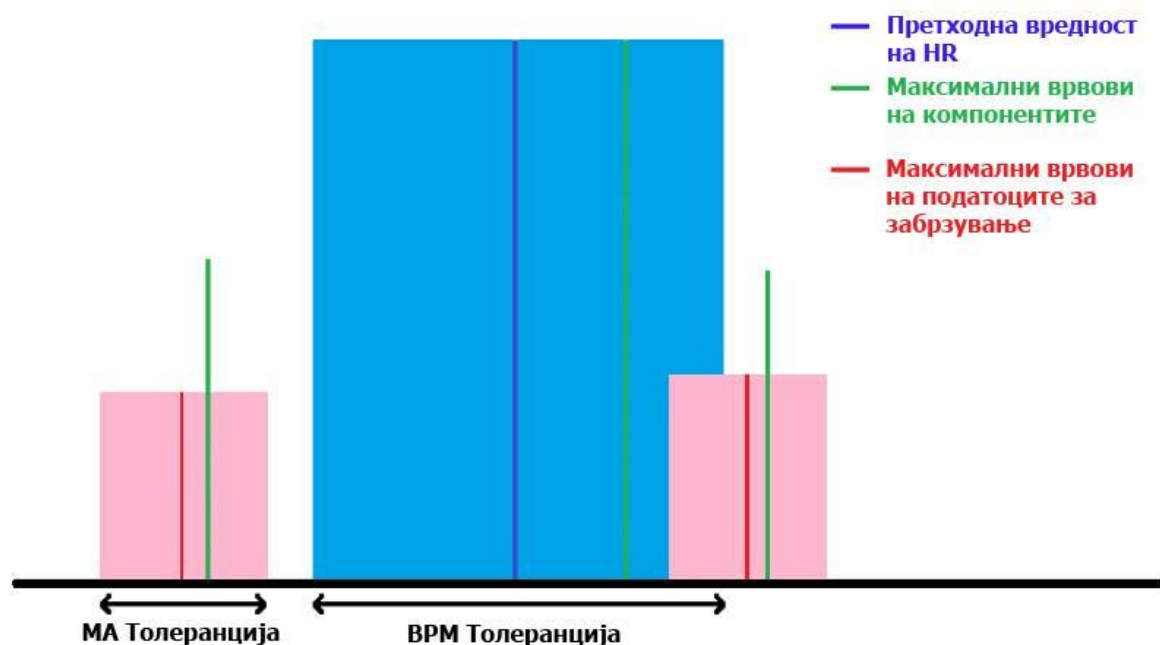
3.4.5. Отстранување на компоненти (RoC)

Овој метод за справување на МА во PPG-сигналот се состои во идентификација на компонентите кои кореспондираат со МА и реконструкција на сигналот без ваквите компоненти, односно се врши отстранување на таквите компоненти. Основната идеја се состои во тоа компонентите кои имаат врв на приближно иста фреквенција како и врвовите на податоците за забрзување (МА компонентите), да се идентификуваат и на тој начин потоа тие да бидат исклучени од реконструкцијата. Конверзијата во фреквентниот домен е направена преку FFT.

3.4.6. Идентификација на МА компоненти

Постојат две можни ситуации за податоците за забрзување или врвовите кореспондираат со МА во PPG-сигналот или во сигналот постои шум од друга природа. Во случајот кога во сигналот постои шум, алгоритмот не бара врвови во податоците за забрзување, бидејќи нема информација за МА. Еден од начините за препознавање дали во сигналот постои шум од друга природа, се прави преку куртоза (kurtosis), што ја отсликува острината на даден врв од сигналот. Ниска куртоза значи дека сигналот е порамен, што значи дека во најголем дел се состои од шум. Следниот чекор е да се лоцираат врвовите на податоци за забрзување кои се одговорни за МА врвовите во PPG-сигналот. Кога амплитудата на врвот во забрзувачките податоци е мала, соодветниот врв во PPG-сигналот исто така ќе биде мал. Поради ова, се земаат предвид, односно, се разгледуваат само врвовите кои имаат амплитуда поголема од 50% од максималниот врв во забрзувачките податоци. Наредниот чекор се состои од споредба на локациите на овие врвови со локацијата на максималниот врв на секоја компонента од PPG-сигналот, добиена од SSA алгоритмот. Поради тоа што врвовите од забрзувачките податоци секогаш не се совпаѓаат со МА врвовите во PPG-сигналот, одредена толеранција мора да биде дефинирана - МА толеранција. Доколку максималниот врв на одредена компонента се наоѓа во рамките на таа толеранција, ваквата компонента се означува како МА компонента. Во случаите пак, кога МА е далеку од вистинскиот HR или пак кога не постои МА, ова решение е доволно. Но, во случајот кога МА се појавува на истата фреквенција или блиску до фреквенцијата на вистинскиот HR, тогаш се јавува проблем, односно, компонентите кои кореспондираат со вистинскиот HR, исто така, ќе бидат исклучени од реконструкцијата. Решение за овој проблем е да се користи информацијата за претходната естимирана вредност за HR: компонентите кои имаат максимален врв околу истата фреквенција како и еден од врвовите на забрзувачките податоци, нема да бидат отстранети доколку се блиску до фреквенцијата на претходно проценетата вредност за HR. Сепак, HR може да се менува значајно меѓу два временски прозорци. Ова значи дека и овде е неопходно да се дефинира одреден вид на толеранција - BMP толеранција. Дополнително,

што може да се направи за засилување на реконструираниот сигнал, е да се отстранат оние компоненти чиј максимален врв е далеку од претходно проценетата вредност за HR. Ваквиот ранг треба да биде доволно голем, така што дури и во случајот кога во проценетата вредност за HR постои голема грешка, сепак, компонентите кои одговараат на HR нема да бидат отстранети. Недостатокот на овој метод се состои во тоа што во голема мера се потпира на фактот дека процената за HR е точна во претходниот временски прозор. Методот нема да функционира доколку претходната проценка за HR има голема грешка. Сепак, овој проблем може да биде решен со следење на спектралните врвови. Ова е илустрирано на слика 3.9.



Слика 3.9: Начин на идентификација на компонентите на движечките артефакти

Figure 3.9 Example of moving artifacts identification

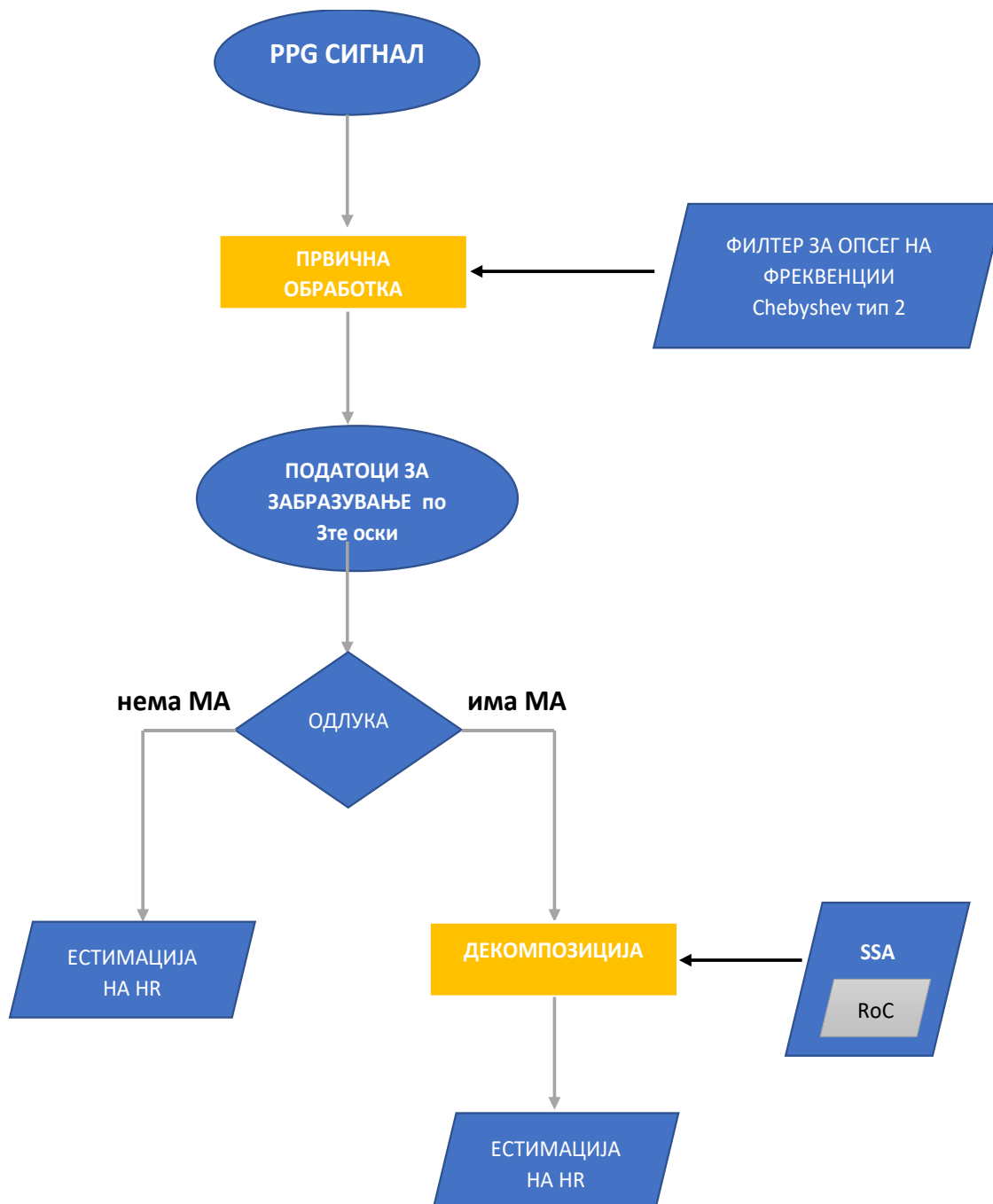
Во овој пример, двете компоненти кои имаат максимален врв во рамките на МА толеранцијата (црвената површина) ќе бидат изоставени од реконструкцијата. Додека пак, компонентите кои имаат максимален врв во рамките на BPM толеранцијата (сината површина) нема да бидат изоставени. Ова ќе бидат компонентите кои го содржат HR сигналот.

Во однос на оптимизацијата и прилагодувањето на параметрите во овој дел, можат да се прилагодуваат 3 параметри за добивање на подобри резултати. Тоа се куртозата, толеранцијата за МА и толеранцијата за претходна БП. За одредување на точниот критериум за куртозата, може да се направи истражување на податоците за забрзување во дадено податочно множество. Така, во случајот кога постојат јасно издвоени врвови во забрзувачите податоци, куртозата на сигналот има вредност поголема од 130. Во врска со МА толеранцијата, оваа вредност треба да биде доволно голема за да ги вклучува компонентите кои одговараат на МА, но во исто време и доволно мала така што не ги вклучува вистинските HR кога МА е близу до таа фреквенција. При анализа на податочните множества, забележано е дека максималната разлика меѓу локациите на врвовите во забрзувачите податоци и соодветните врвови во PPG-сигналот е околу 7 BPM. Затоа се дозволува толеранција од ± 7 BPM, која би требало да биде доволна за одредување на компонентите кои припаѓаат на МА. За проценување на точна вредност за BPM толеранција, потребно е да се одреди максималната промена во HR помеѓу два последователни временски прозорци. За да бидеме сигурни дека нема да биде отстранета компонента која припаѓа на HR, избрана е толеранција од ± 15 BPM.

Во нашиот алгоритам, пред да почнеме со делот за декомпозиција на сигналот, користиме други два дополнителни чекора. Првиот дел е првичната обработка на сигналот (претпроцесирање). Во овој дел за разлика од околината TROIKA, ние користиме филтер за опсег на фреквенции (bandpass filter), конкретно филтерот Chebyshev тип 2. Овој филтер според (Liang et al., 2018) се покажува како оптимален филтер за ваков тип на сигнали, од сите други 9 испитувани различни филтри. Испитувањата се правени на 3 различни множества од податоци поделени според различниот квалитет на сигналот (сигнал што има правилна форма, сигнал што ја нема задржано правилната форма, но сè уште добро се разликуваат врвовите и сигнал кој целосно е зафатен со шум). Од овие причини во делот за првична обработка ние донесовме одлука да го користиме овој тип на филтер.

Вториот чекор кој го додаваме за да се намали времето на обработка на сигналот е делот со правење на одлука од две можности, како што се прави и во (Song et al., 2017). Оваа одлука се прави за да се види дали има потреба воопшто да се оди во делот на декомпозиција или директно да се премине кон делот за естимација на HR. Овој чекор има за цел, преку амплитудата од периодограмот на податоците за забрзување, да се одлучи дали постојат МА или не. МА постојат кога спектралната амплитуда од секоја оска на забрзувањата е поголема од 0.1 (густина на спектарот на моќност - power spectral density). Во секој случај постојат други алгоритми кои го определуваат присуството на МА, како што е коефициентот на корелација (Yangsong Zhang et al., 2015).

Дури откако ќе поминат овие два чекори се оди во делот за декомпозиција на сигналот со помош на SSA алгоритмот во кој е вклучен делот за отстранување на компонентите RoC. Графичкиот приказ на нашиот алгоритам е даден на слика 3.10.



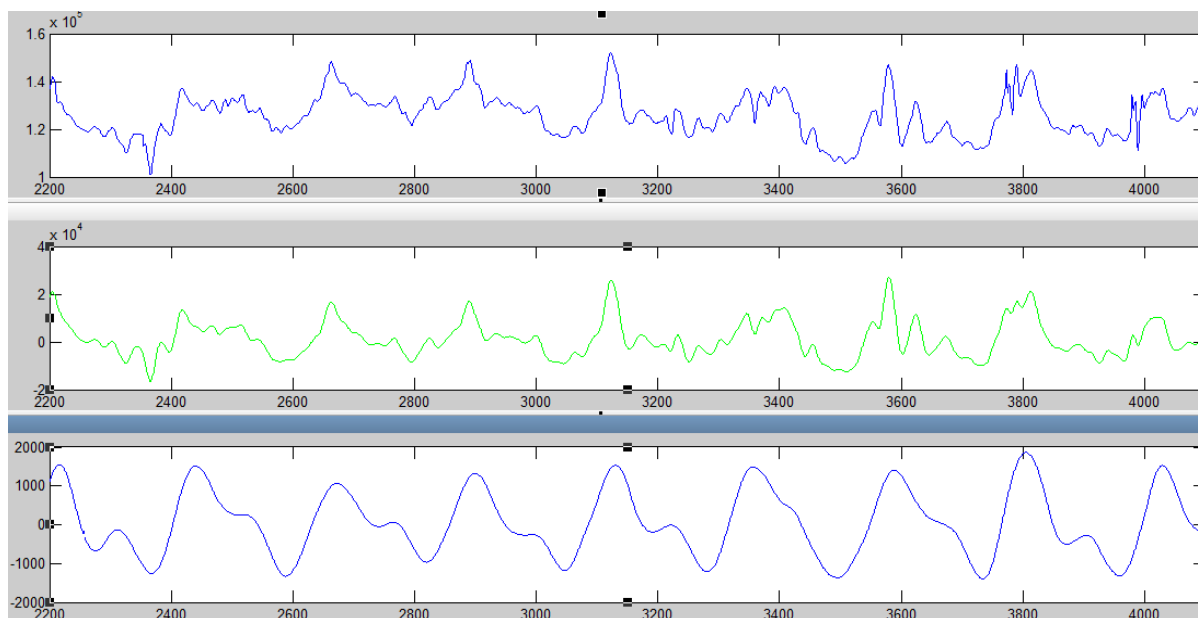
Слика 3.10: Шематски приказ на алгоритамот

Figure 3.10 Schematic representation of the algorithm

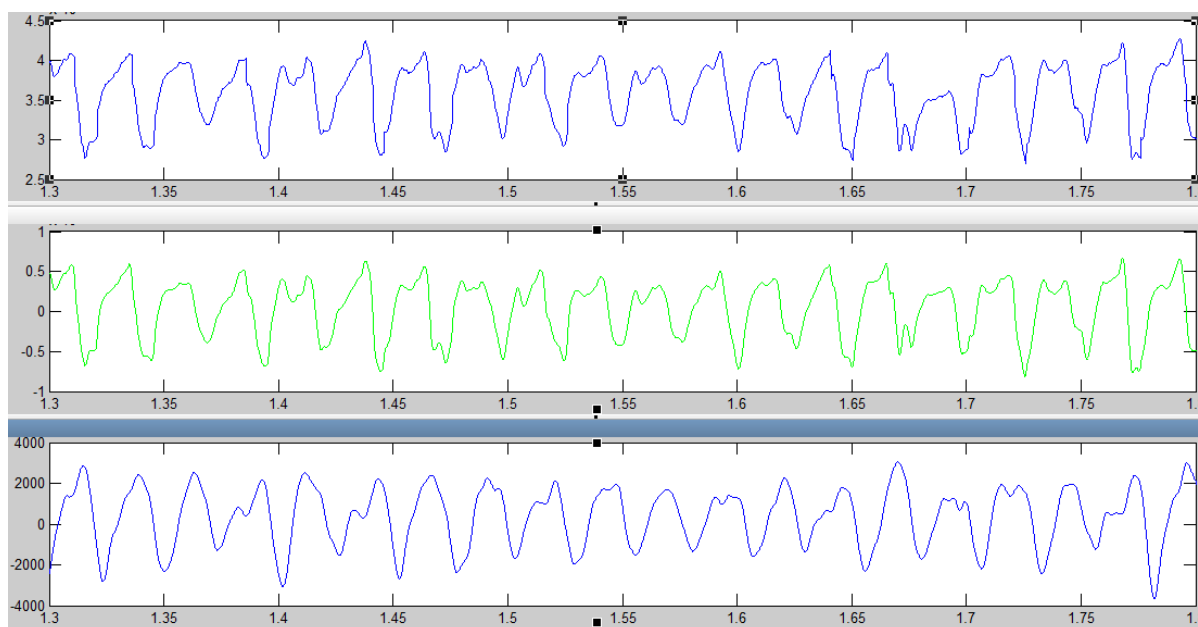
3.4.7. Резултати од применетиот алгоритам

Алгоритамот со соодветните штелувања и оптимизации го применивме на нашето множество на податоци. Како што веќе спомнавме, мерењата се направени со сензор поставен на зглобот од раката. Уредот работи со фреквенција од 200 Hz.

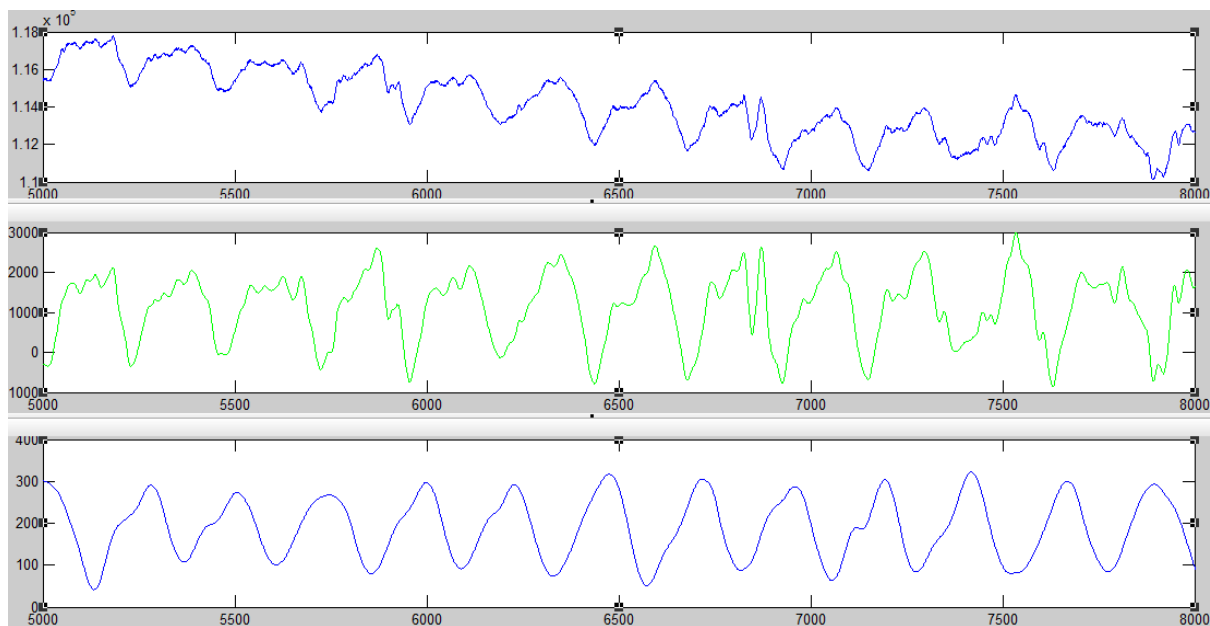
Алгоритмот за филтрирање го тестиравме на 15 различни мерења. Дел од резултатите од филтрирањето, поточно 3 мерења се прикажани во продолжение (слика 3.11 а), б) и в)) Првиот сигнал кај секое од мерењата е почетниот, односно измерениот PPG сигнал, вториот дел е сигнал добиен од првичното филтрирање (во фазата за претпроцесирање со Chebyshev 2 од тип 2 филтерот), а третиот е крајно филтрираниот сигнал со отстранетите движечки артефакти.



а) Прво мерење



б) Второ мерење



в) Трето мерење

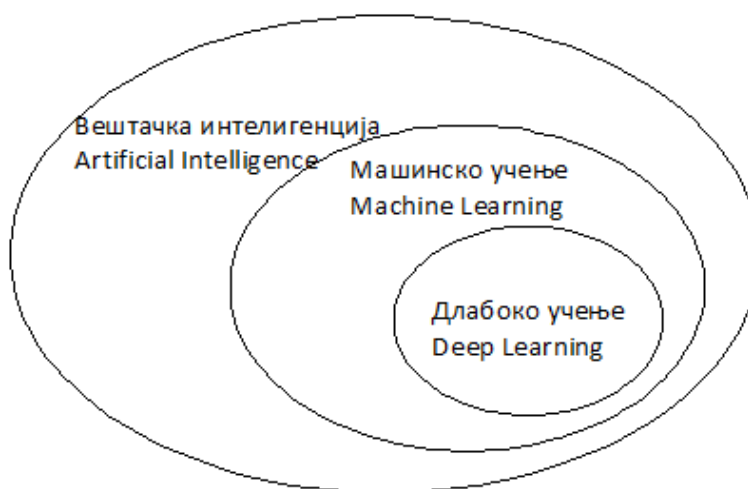
Слика 3.11: Сегменти од резултатите од филтрирањето применети на 3 различни мерења. Првиот дел кај секој од сигналите е измерениот PPG сигнал, вториот е сигнал добиен од првичното филтрирање и третиот е крајно филтрираниот сигнал со отстранетите движечки артефакти.

Figure 3.11 Segments from filtering results applied to 3 different measurements. First part of each measurement is measured PPG signal, second part is signal obtained from the initial filtering and third part is filtered signal with the removed moving artefacts

4. Методи на машинско учење

Машинско учење (Machine Learning - ML) претставува алгоритамски пристап, за решавање на секоја дадена задача, што дозволува можност компјутерот да извршува задачи без експлицитно програмирање. Ваквиот пристап овозможува човечка интелигенција и можност за учење преку искуства и се развива на тој начин што во секоја итерација се обидува да се доближи до оптималното решение. Затоа се вели дека системот за машинско учење се обучува, наместо експлицитно да се програмира. Тој е претставен со многу примери релевантни за задачата и наоѓа статистичка структура во овие примери што на крајот му овозможува на системот да излезе со правила за автоматизирање на задачата. Методите на машинското учење можат да се искористат и во процесот на естимација на крвниот притисок од PPG сигналите, како и да помогнат понатаму во процесот на предвидување (Witten et al., 2016).

Машинското учење може да се дефинира како гранка од вештачката интелигенција и научна дисциплина која се занимава со дизајнирање и создавање на алгоритми кои овозможуваат компјутерските системи да ја подобруваат својата работа со помош на емпириски податоци, т.е. податоци добиени преку механички сензори или база на податоци или податоци добиени пат на експерименти и набљудување.



Слика 4.1: Вештачка интелигенција, машинско учење и длабоко учење

Figure 4.1 Artificial intelligence, machine learning and deep learning

Машинското учење може да се подели во два главни дела и тоа, надгледувано (supervised learning) машинско учење и ненадгледувано (unsupervised learning) машинско учење, но, исто така, техниките на машинското учење можат да се поделат и во повеќе дополнителни категории во зависност од нивните карактеристики како што се делумно надгледуваното (semi-supervised learning) учење и учењето со поттикнување (reinforcement learning).

- **Надгледувано учење**

Надгледуваното учење е алгоритам за наоѓање на функција од означените податоци и тренираните податоци. Во надгледуваното учење, податоците за тренирање се состојат од влезни вектори и посакуван излезен одговор, кој исто така е опишан како надгледуван сигнал. Алгоритамот ги анализира обележаните податоци и генерира функциски блок на заклучување за да обработи ново множество на податоци. Идеално сценарио би било кога би се овозможило алгоритамот да се тренира и потоа прецизно да се изведе класификацијата за новите влезни примери на податоци.

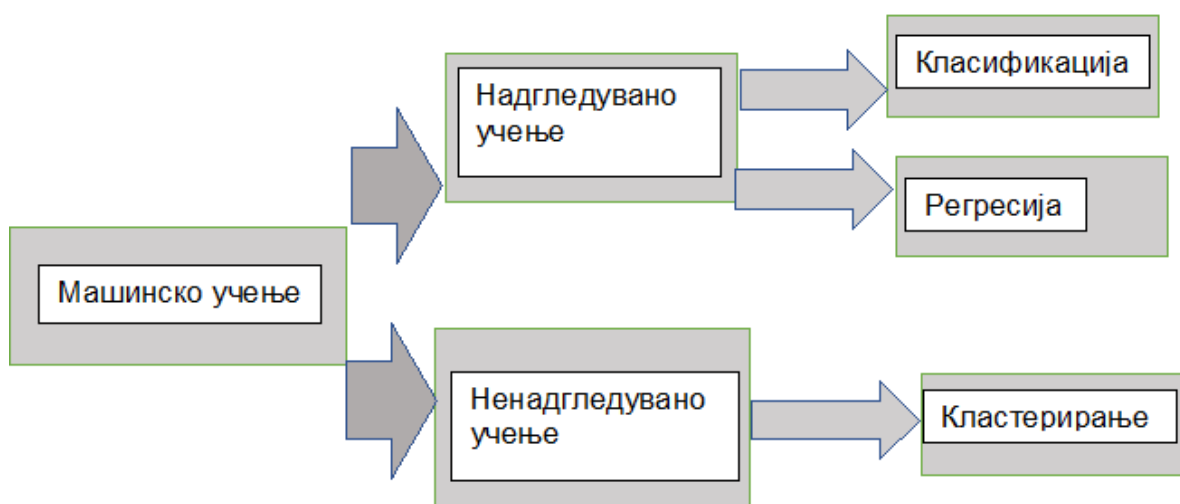
Значи, кај оваа категорија од методите на машинското учење, моделот се полни со група од влезни множества и соодветни излези. Целта е да се најде функцијата на мапирање која ги поврзува влезните променливи и дадените излези.

Надгледуваното учење понатаму може да се подели на модели на класификација и модели на регресија. Класификацијата се користи кога излезните променливи можат да се поделат во категории, на пример, „хипертензија со оцена 1“, „хипертензија со оцена 2“, „нормален крвен притисок“ и сл. Односно, класификацијата претставува процес во кој излезот може да припаѓа во дадена класа. А, регресијата е процес во кој излезните податоци се реални вредности. Односно, моделите на регресија се користат кога излезите се континуирани променливи, како што е вредноста за крвен притисок.

- **Ненадгледувано учење**

Машинското учење без надгледување се користи во случај кога податоците се неозначени и неклассифицирани. На алгоритмот му се дава множество податоци кое го обработува, за да се дефинира скриената структура и карактеристичните разлики. Тогаш, податоците се групираат според сличноста на карактеристиките. Бидејќи на алгоритмот му се дадени податоци без ознака, не може да се изведе евалуација на точноста на излезот. Значи, за разлика од надгледуваното во оваа категорија од машинското учење, на моделот му е дадено множество од карактеристики, кои немаат ознаки, за податоците да се поделат во кластери, земајќи ги предвид нивните сличности. Главната задача во учењето без надзор е: кластерирањето и асоцијацијата.

- Кластерирањето е процес на добивање на наследни групи кај податоците, со други зборови, групирање на податоците според сличноста на нивните карактеристики.
- Асоцијацијата е процес на наоѓање на правила преку кои податоците се групираат во делови какви што се кластерите.



Слика 4.2: Поделба на машинското учење

Figure 4.2 Parts of machine learning

Дополнителни категории на машинското учење кои ги спомнавме погоре можат да бидат: делумно надгледувано учење и учење со поттикнување (Reinforcement learning)

- **Делумно надгледувано учење**

Ваквиот вид на учење е мешавина од надгледуваното и ненадгледуваното учење, каде дел од податоците не се означени. Ова учење ги комбинира надгледуваното и ненадгледуваното учење и најчесто во тренинг примерот се вклучени податоци добиени од човекот, но и податоци добиени со податочно рударење.

Преку овој пристап, означените и неозначените податоци можат да бидат пуштени низ систем за кластерирање (ненадгледувано учење), преку кое се означуваат неозначените податоци врз основа на сличностите, а потоа се користат резултантните податоци да се направат предвидувања на новите податоци (надгледувано учење)

- **Учење со поттикнување (Reinforcement learning)**

Кај оваа категорија на машинско учење, за да се обезбеди решение за специфичен проблем, се користат пробни и грешни пристапи. При решавањето се изведува случајна акција и се добива награда доколку резултатот придонесе за постигнување на целта или казни ако се случи спротивното. Ова учење дава делумна автономност во делувањето со тоа што претходно се врши набљудување и проучување на околината. Секоја акција има некоја последица по околината која зависно од акцијата дава позитивна, негативна повратна врска или помош до машината која учи така што таа може да ги контролира своите наредни акции.

Кога станува збор за естимација на крвниот притисок, фокусот главно се става на методите за надгледувано учење, а во одредени случаи и методи за ненадгледувано учење, бидејќи целта е да се предвидат вредностите на крвниот притисок, што е континуирана променлива. Затоа ние во ова поглавје се задржуваме на основните методи од машинското учење кои се користат за естимација, класификација и кластерирање на крвниот притисок.

4.1 Кластерирање

Кластерирањето претставува процес на групирање на податоците во класи или кластери врз основа на нивните сличности. Тоа значи дека објектите во даден кластер се многу слични помеѓу себе, но различни од објектите од другите кластери.

Кластер анализата може да се користи како алатка за податочно рударење за да се разберат карактеристиките на секој кластер, може да се користи и како пред процесирачки чекор кај други алгоритми (избирање на атрибути, карактеризација) или да се фокусира на одредено множество на кластери за нивна понатамошна анализа.

4.2 Методот K-means

За дадено податочно множество од n објекти, а k бројот на кластери, алгоритмот на делење ги организира објектите во k дела (кластери). K-means е најкористениот метод за делба, односно кластерирање кои се користи и во проблематики како што е нашата, затоа овде на кратко го објаснуваме овој метод. Методот на K-means го дели множеството од n објекти во k кластери, а при тоа резултатот да биде висока сличност внатре во кластерот а ниска сличност помеѓу кластерите.

Кластерот е средна вредност од објектите во кластерот, кои го карактеризираат секој кластер.

Алгоритмот k-means продолжува на следниов начин: случајно избира k објекти како почетни центри на кластери. Центрите на кластерите се наречени центроиди. Секој објект кој е избран по случаен избор, може да биде почетен центар. Откако ќе се изберат центроидите се одредуваат централните координати и се определува растојанието од секој објект до центарот. Понатаму, се врши групирање на објектите според големината на минималното растојание. Според погоре кажаното, алгоритмот може да се претстави во следните чекори:

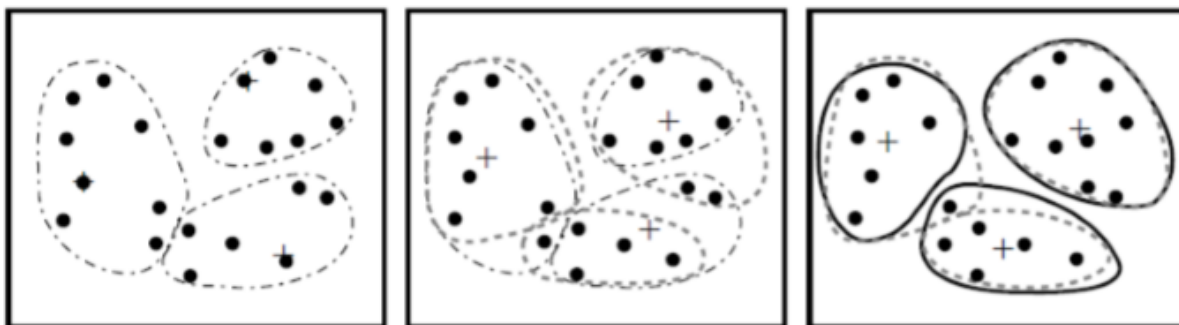
1. Се избираат K-точки во просторот претставен од објекти кои е потребно да се кластерираат. Избраните точки ја сочинуваат иницијалната група на центроиди.

2. Се врши доделување на секој од објектите во групата која го поседува најблискиот центроид.
3. Кога секој од објектите е сместен во соодветна група-кластер, се врши повторно пресметување на позициите на центроидите.
4. Се повторуваат чекорите 2 и 3 се додека има промена во позицијата на центроидите.

Критериумот за квадратна грешка се користи за пресметување на квадратната грешка за сите објекти во податочното множество, а се дефинираа на следниот начин:

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (4.1)$$

каде p е точката која го претставува објектот а m_i е средината на кластерот C_i . Алгоритамот се обидува да определи k делови кои ја минимизираат функцијата на квадратна грешка (Han et al., 2012). На слика 4.3 е престапен дел од процесот за кластерирање.



Слика 4.3: Постапкана кластерирање

Figure 4.3 Clustering procedure

Иако процедурата на k-means секогаш завршува, односно конвергира, не секогаш овој алгоритам ја наоѓа најоптималната конфигурација. Освен тоа, алгоритамот е значајно чувствителен на почетниот случаен избор на центроидите. За да се намали овој ефект, k – means може да се изврши неколку пати, а потоа да се направи споредба на добиените резултати во секој од обидите и да се избере најдобриот.

Дополнителна слабост на овој алгоритам е тоа што резултатот зависи од вредноста на бројот на кластери – k . Сепак не постои генерално теоретско решение за наоѓање на оптималниот број на кластери за определено податочно множество. Иако постојат неколку методи кои се познати и користени за одредување на оптималниот број на кластери, сепак наједноставно решение во овој случај, е да се направат неколку обиди со различни вредности за бројот на кластери и да се изврши споредба на добиените кластери во секој од направените обиди.

4.3 Класификација и регресија

Бидејќи целта е да се естимира BP, што не е дискретна туку континуирана променлива, потребно е справување со проблем на регресија. Во анализата со регресија се прави обид да се добие функциска врска помеѓу континуираната зависна променлива (BP) и независните променливи (PPG карактеристики). Таквата врска може да биде едноставна (на пр. линеарна) или комплексна (на пр. нелинеарна). Врз основа на векторот на карактеристиките на PPG-сигналот, се користат различни методи за машинско учење за естимирање на систолниот (SBP) и дијастолниот (DBP) притисок.

4.3.1 Класична регресија

Регресивните модели вклучуваат:

- Непознатите параметри на моделот, означени со α .
- Независните променливи означени со X .
- Зависните променливи означени со Y .

Алгоритамот ја наоѓа функцијата f која ги поврзува X и Y што е можно подобро, по даден конкретен критериум. Ова може да се формулира како:

$$Y \approx f(X, \alpha) \quad (4.2)$$

Во случај на естимација на крвниот притисок од PPG сигнали, векторот X ги содржи пресметаните карактеристики, кои го опишуваат PPG-сигналот. Зависната променлива Y може да биде или систолниот (SBP) или дијастолниот (DBP) притисок, а параметрите α зависат од конкретниот алгоритам. Постојат повеќе

регресивни модели кои се применуваат за ваквата проблематика. Тие се опишани во продолжение. Во голем број случаи регресивните и класификациските модели се многу слични.

Алгоритмите за регресија можат да бидат едноставни како моделот за линеарна регресија или посложени како што се невронските мрежи. Не постои некоја генерална формула за сите проблеми, бидејќи секој од моделите има свои предности и недостатоци и треба да се најде некоја рамнотежа за да се дефинира најдобар модел.

4.3.2 Линеарна регресија (linear regression)

Линеарната регресија претставува линеарен пристап за моделирање на функцијата f . Оваа регресија се обидува да ги моделира дадените тренирачки податочни точки со линеарна функција, која одговара на поставување линија на даденото множество на познати податочни точки. Линијата се поставува со користење на пристап на најмали квадрати. Се решава преопределен збир на линеарни равенки, со употреба на проценки на решенија кои произведуваат најниска сума на квадратни грешки (sum of squared errors -SSE).

4.3.2.1 Регуларна линеарна регресија (Regularized Linear Regression -RLR)

Линеарните регресивни модели кои се соодветно регулирани со К-пати вкрстена валидација, се обучени за проценка на линеарноста на проблемот. Доколку карактеристичниот вектор и целта имаат силна нелинеарна врска во тој случај ваквиот начин не може да се примени. Во секој случај, ваквите модели се едноставни, лесни за обука и помалку склони на преклопувања во споредба со другите можни алтернативи. Исто така, се потребни помалку примероци за обука и затоа се поефикасни, а со тоа и нивната имплементација е поефикасна.

Како што спомнавме, линеарната регресија е најпроста форма на регресија. Целта на ваквата регресија е да се моделираат одредени тренирани податоци со помош на линеарна функција. На овој начин предвидената вредност се добива како резултат на линеарна комбинација од карактеристиките.

$$\hat{y}(\omega, x) = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n \quad (4.3)$$

каде \hat{y} е предвидената вредност, $X = (x_0, x_1, x_1, \dots, x_n)$ се пресметаните карактеристики, $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ е векторот со коефициенти и ω_0 е вредноста на прекин (interceptvalue). Целта е да се најдат ω и ω_0 кои подобро се вклопуваат во тренирачките податоци и потоа да се направи предвидување на сите нови X . Процесот на наоѓање на најдобра вредност за ω , може да се постигне со минимизирање на функцијата за трошок J , како обични минимални квадрати.

$$\min(J(\hat{Y}, Y)) = \min\left(\frac{1}{n} \sum_1^n (\hat{y} - y)^2\right) \quad (4.4)$$

4.3.2.2. Ридж линеарна регресија (Ridge Linear Regression - RLR)

RLR е исто така вид на линеарна регресија која може да се користи во естимација систолниот и (SBP) и дијастолниот (DBP) притисок, а исто така е погодна за случајот кога карактеристичниот вектор и целта имаат силен линеарен сооднос. Таа има за цел да ја минимизира функцијата на загуба, дадена на следниот начин:

$$\min_{\omega} \left\| X\omega - y \right\|_2^2 + \alpha \left\| \omega \right\|_2^2 \quad (4.5)$$

Каде X, ω, y и $\alpha \geq 0$ се карактеристичниот вектор, коефициентот, целта и параметарот на сложеноста, соодветно. Естиматорот се пресметува со следната равенка откако ќе се научи коефициентот, каде X_i, ω_i, X_i , се компоненти на X и ω , соодветно, а N е должината на X .

$$y = \sum_{k=1}^N \omega_i X_i \quad (4.6)$$

4.3.2.3 Повеќекратна линеарна регресија (Multiple Linear Regression - MLR)

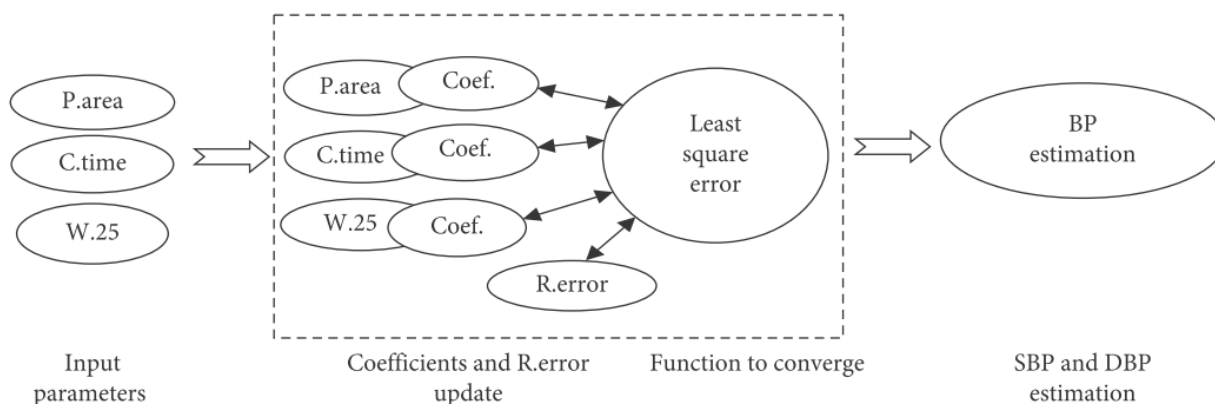
MLR е вид на алгоритам за машинско учење кој е широко користен за проценка на крвниот притисок (Buxi et al., 2015) (Gesche et al., 2012) (R. Wang et al., 2014) Алгоритамот започнува со случаен избор на коефициенти од линеарниот алгоритам $(\theta_0, \theta_1, \theta_2, \theta_3)$. Секој предиктор-предвидувач е поврзан со одреден

коэффициент како што е прикажано на Слика 4.4. По секоја итерација, коефициентите и случајната грешка (ε , разликата меѓу проценетата и референтната вредност за крвниот притисок) се ажурираат. За минимизација на квадратната грешка – како што е прикажано во следното равенство, се користи алгоритмот за најмал квадрат. Итеративната минимизација на квадратната грешка продолжува се до конвергенција кога се генерира проценката за вредноста на крвниот притисок:

$$J(\theta_0, \theta_1, \theta_2, \theta_3) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (4.7)$$

$$h_{\theta}(x) = (\theta_0 + \theta_1(\text{area}) + \theta_2(\text{crest}_{\text{time}}) + \theta_3(\text{Width}_{25})) + \varepsilon \quad (4.8)$$

Каде m се вкупните податоци за тренинг (обучување) ε е случајна грешка, θ_{0-3} се коефициенти, $h_{(x)}$ = проценетата вредност за крвен притисок и y = референтната вредност за крвниот притисок (Khalid et al., 2018).



Слика 4.4: Дијаграм на тек на MLR (Khalid et al., 2018)

Figure 4.4 MLR flow diagram (Khalid et al., 2018)

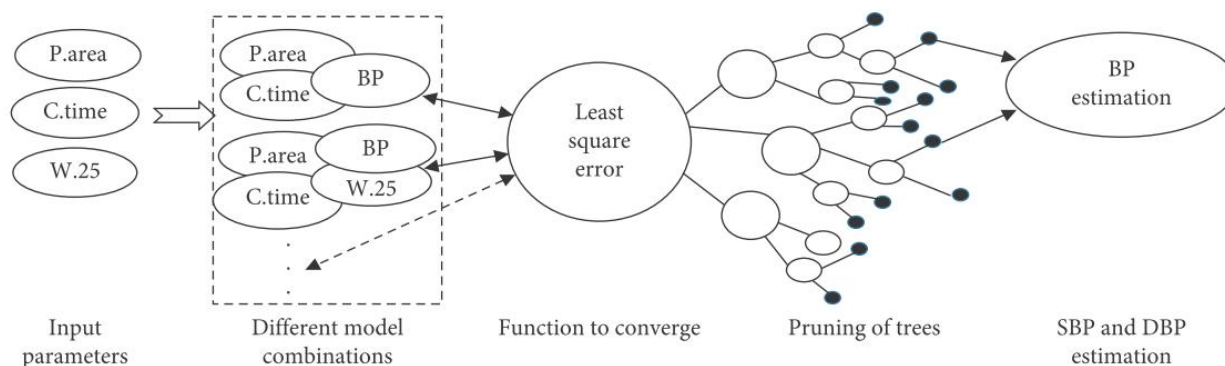
4.3.3 Регресија и класификација на основа на дрва (Regression trees)

Регресивни дрва претставува нелинеарен пристап, бидејќи тие прво рекурзивно го делат просторот на податочните точки во помали региони, а потоа сите се вклопуваат во едноставен константен модел (на пр., средина, медијана и сл.) во секоја од партициите. Поделбата на податоците во смисла на минимизирање на сумата на квадратните грешки во сите потенцијални делови е пресметковно неостварлива, така што најчесто се користи алчен метод за да се оствари тоа.

Регресивните дрва претставуваат друг непараметарски пристап на машинско учење за предвидување. Тоа е релативно брз алгоритам за тренирање на податоците. Ваквиот алгоритам ги носи одлуките од коренските јазли до јазлите – листови. Регресивните дрва се бинарни дрва и јазолот - лист што ги содржи одговорите, е во нумеричка форма. Податоците, овде, се делат со најдобриот критериум за оптимизација. Критериумот за стопирање на поделбата за да се направи чист јазол, врз основа на средна квадратна грешка (MSE) е следниот:

$$\text{MSE (надгледуван одговор)} < \text{MSE (надгледуван одговор од вкупните податоци)} \times \text{толеранција} \quad (4.9)$$

Чист јазол покажува дека средната квадратна грешка на надгледуваните одговори е помала од средната квадратна грешка од сите податоци помножена по вредноста за толеранција. За оптимизација, алгоритмот ги дели гранките од дрвата за да ја минимизира грешката на предвидување, како што е прикажано на слика 4.5.



Слика 4.5: Дијаграм на тек кај регресивни дрва (Khalid et al., 2018)

Figure 4.5 Flow diagram of regression trees (Khalid et al., 2018)

Регресија на дрвото на одлуки (Decision Tree Regression - DTR) е воведена во (Leo Breiman et al., 2017). DTR има структура како дрво, која е составена од многу јазли. Јазолот на листот ја одредува вредноста на предвидувањето. Точноста на DTR е генерално конкурентна со линеарната регресија. Дава подобри резултати во случај кога карактеристичниот вектор и целта имаат нелинеарна структура, но од друга страна има тенденција да биде нешто помалку точен за проблемите во кои постои добра линеарна врска.

4.3.3.1 Модел дрва (*Model trees*)

Моделот дрва се слични со регресивните со таа забележителна разлика што, вградениот модел во секоја од партициите е функција (на пр., линеарна функција) а не е константна вредност како кај регресивните дрва.

4.3.3.2 Ансамбли на дрва (*Ensembles of trees*)

Во одредени случаи дрвата можат да се спојат во ансамбли на дрвја, кои често покажуваат супериорни перформанси во споредба со одделни дрвја. Ансамблот се заснова на агрегација на bootstrap (bagging - пакување) или засилување (зголемување - boosting). Пакувањето, многу пати ги вклопува дрвата во различни bootstrap примероци на податоци, на тој начин, секој пат, создавајќи малку поразлични дрва. Предвидувањето на пакувањето е просечното предвидување од овие дрва. Зголемувањето односно засилувањето, од друга страна, создава предодредено (default) дрво на сите податоци и потоа создава дополнителни дрва фокусирајќи се на податоците што предодреденото дрво ги има предвидено со најголема грешка. Крајното предвидување е дадено како тежински просек на индивидуалните предвидувања на дрвото.

4.3.3.3 Регресија на пакување (*Bagging Regression - BR*)

Регресијата на пакување е метод за ансамбл за надгледувани регресивни проблеми (L Breiman, 1996). Таа гради модел базиран на основен ученик. За секој основен ученик, примероците за учење се извлекуваат случајно со замена од оригиналното множество на податоци. Естиматорот на ваквата регресија се добива со земање на просечната вредност од сите нивни индивидуални предвидувања. Таквиот ансамбл естиматор обично може да дејствува како начин да се намали варијацијата на основниот естиматор. Во (Xie et al., 2019) се користи DTR како основен ученик.

4.3.3.4 Случајни шуми (*Random Forest*)

Случајните шуми се слични како и ансамблиите, со забележителна разлика што алгоритмот секој пат креира дрва користејќи различни множества на карактеристики за разделување. На овој начин се создаваат многу по различни

дрва кои не се многу во корелација едни со други. Крајното предвидување и овде е дадено како просек од индивидуалните предвидувања на дрвата.

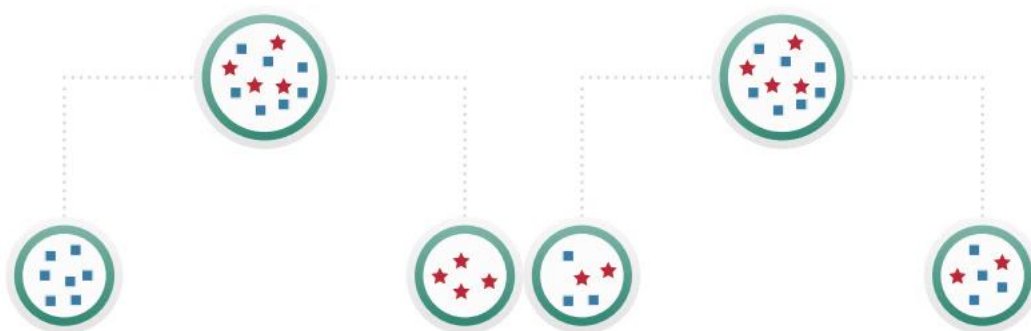
Случајна шума (random forest-RF) е пристап кој може да се користи за класификација и регресија. Тој претставува методот на учење со ансамбл и користи неколку алгоритми за предвидување за да се подобри целокупната точност на системот (James et al., 2013).

Случајна шума за регресија се користи за да се направи регресивна анализа која како таква е претставена во (L. Breiman, 2001). Тоа е учење во ансамбл што гради модел со DTR како основен ученик. Секој основен ученик е трениран за случајно подмножество на тренирачките примероци. Случајниот избор на карактеристики е дополнително воведено за време на постапката за обука на DTR. Крајното предвидување се добива со земање на средна вредност од сите дрва.

Овој метод користи дрва на одлуки кои претставуваат основни структури. Дрво на одлука претставува едноставен алгоритам инспириран од самиот процес на одлучување кај луѓето, и заради овој факт обично е добро разбран од корисниците за разлика од алгоритмите на црната кутија, како што се невронските мрежи NN. Дрвото се гради од врвот па до дното, каде што секој јазол претставува одредена карактеристика. Овој процес може да се поедностави во неколку чекори:

1. Избор на податоци за тренирање;
2. Избор на најдобри карактеристики;
3. Поделување на податоците според можните вредности на избраната карактеристика (создавање јазол);
4. Рекурзивно повторување на чекорите 2 и 3 со подмножество на податоци што произлегуваат од последниот чекор;
5. Стопирање кога ќе се достигне критериумот за запирање.

Изборот на карактеристики може да се направи користејќи неколку критериуми, како што се мерка на растојание (distance measure), рата на добивка (gain ratio) или критериуми кои се базираат на загаденоста на карактеристиките (impurity based). Основната цел е да се избере карактеристика што подобро го опишува проблемот (слика 4.6).



Слика 4.6: Разлика во добивката на информации во зависност од избраните карактеристики. Во првиот случај има голема добивка на информации, а во вториот мала

Figure 4.6 Difference of gain information depending on selected features. In the first case there is a large gain of information and in the second small

Големината на дрвото, исто така, треба да се контролира бидејќи поголемо дрво е повеќе склоно на преклопување и постапката на тренирање трае подолго. Затоа можат да се користат неколку услови како максималната длабочина на дрвото и максимален број на лисја. Може да се направи дополнителен чекор на поткастрување, така што ќе се отстранат непотребните поделби. Овој процес го поедноставува дрвото и води кон под-дрво што има помалку преоптоварување и преклопување на тренирачките податоци.

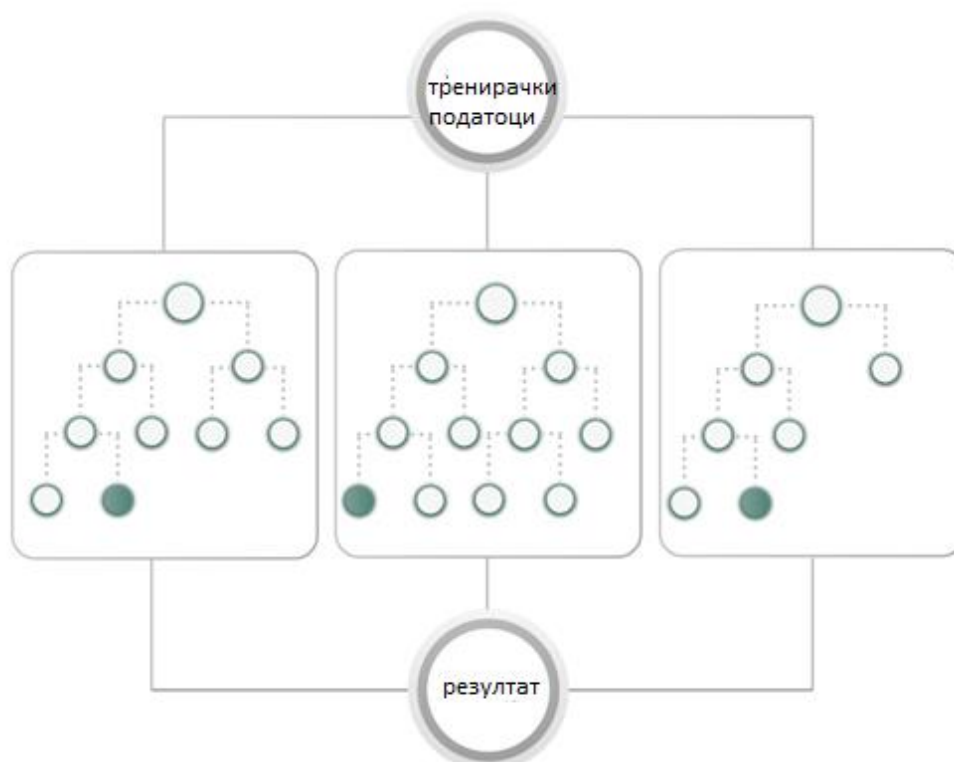
Дрвата немаат толку високо ниво на прецизност за разлика од други методи на класификација и регресија. Како решение за да се подобри преведувачката способност на дрвата е со користење на случајни шуми.

Подобрувањето може да се објасни со користење на статистички концепти. Ако е дадено множество од n независни набљудувања $F = \{F_1, F_2, \dots, F_n\}$, секое со варијанса σ^2 , тогаш средната варијанса на набљудувањата е еднаква со средната на варијансите.

$$V(\bar{F}) = \frac{1}{n} \sum_{i=1}^n V(F_i) = \frac{\sigma^2}{n} \quad (4.10)$$

Единственото прилагодување кое е неопходно да се направи е засилувањето (bootstrap - земање на повторувачки примероци од податочното множество), бидејќи е достапно само едно податочно множество. На овој начин со

користење на случајни шуми се создаваат неколку дрва на одлучување. Но, за разлика од методот на пакување (Bagging method), каде дрвата се креираат со избирање на најдобрата карактеристика од секој јазол од сите можни карактеристики, овде избраната карактеристика е од група со определен број на случајно избрани карактеристики. Оваа промена ја подобрува декорелацијата меѓу дрвата.



Слика 4.7: Претставување на методот за случајни шуми

Figure 4.7 Presentation of the random forest method

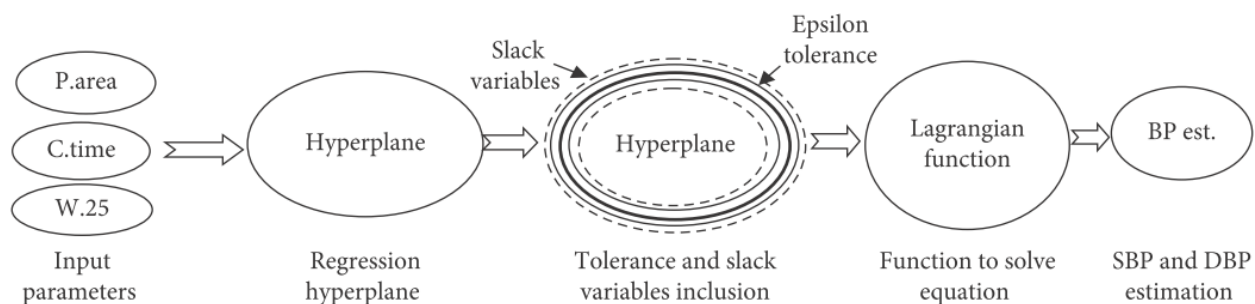
Постојат и други познати методи на регресија како што се машини со носечки вектори (Support Vector Machines - SVM).

4.3.4 Машини со носечки вектори (Support vector machine)

Машини со носечки вектори се множество на поврзани надгледувани методи за учење кои користат класификација и регресија. За дадено множество од тренинг примери, секој од примерите се маркира како припадник на една од две категории. Со помош на векторско-потпорните машини се прави модел со кој се предвидува дали даден влезен пример припаѓа во една или друга категорија. Машините со

носечки вектори се меѓу најмоќните алгоритми за учење во однос на создавање силни модели со разумен напор за тренирање и висока толеранција на шум.

SVM е непараметарски алгоритам кој користи кернел (јадро) функција. SVM-регресијата има слична цел како и во методот на најмал квадрат – MLR, а тоа е минимизација на грешката (квадратна грешка меѓу проценетата и референтната вредност за крвниот притисок). Сепак, пристапот за минимизирање е различен од оној на MLR и се состои во тоа што го користи епсилон (ε) и целта е да се најде функција чија грешка не е поголема од ε . ε -SVM има две граници низ линијата на регресија како што е прикажано на слика 4.8. Сепак, во реалноста, не сите остатоци се наоѓаат во епсилон границата. Поради ова, се воведени нови слаби променливи - друга граница, за да се покријат сите преостанати остатоци, како што е прикажано со испрекинатата линија низ регресивната линија на слика 4.8. Слабите променливите се воведени за да се постигне двојна цел (dual objective). При секоја итерација се ажурираат векторите кои постојат во двојната цел, и равенката се решена аналитички со функцијата на Лагранж.



Слика 4.8: Дијаграми на тек кај методот на SVM (Khalid et al., 2018)

Figure 4.8 Flow diagrams of the SVM method (Khalid et al., 2018)

Кај SVM, критериумот за конвергенција се базира на следново равенство:

$$\Delta = \frac{J(\beta) + L(\alpha)}{J(\beta) + 1} \quad (4.11)$$

каде што $J(\beta)$ е наречена примарна цел. $L(\alpha)$ е двојна цел која се решава со функцијата на Лагранж. Целта е да се минимизира оваа функција за да се добијат естимираните вредности за крвниот притисок. Δ го претставува физибилити јазот (feasibility gap) За конвергенција на алгоритмот, физибилити јазот треба да биде помал во споредба со јазот во рамките на толеранција.

Врз крвниот притисок влијаат различни фактори, а врските помеѓу карактеристиките и крвниот притисок се комплицирани и нема јасно дефинирани механизми. За да се прилагодат податочните множества на нелинеарноста која најчесто јавува и да се надминат недостатоците на традиционалните методи за подесување се користи (support vector regression - SVR) која е искористена за конструирање модели за оценка на SBP и DBP (S. Chen et al., 2019).

SVR има многу предности при решавање на проблеми кои се состојат од мал број на примероци и проблеми кои имаат нелинеарни регресији. Во SVR, влезниот примерок \mathcal{X} е мапиран во високо димензионален карактеристичен простор со помош на нелинеарно мапирање $\Phi(x)$, а потоа е поставен линеарен модел во карактеристичниот простор за да се пресмета функцијата на регресија. Се користи следната равенка:

$$f(x, y) = \omega \cdot \Phi(x) + b, \quad (4.12)$$

во која ω тежински вектор, а b е прагот. Најчесто се користи ε -insensitive функција на загуба, а соодветниот SVR се нарекува ε -SVR. Потребно е да се решат следните ограничувачки проблеми при оптимизацијата:

$$\min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^l (\xi_i + \xi_i^*), \quad \text{така што} \quad \begin{cases} y_i - (\omega^T x_i + b) < \varepsilon + \xi_i \\ (\omega^T x_i + b) - y_i < \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}, \quad (4.13)$$

каде c е фактор на казна, а ξ_i и ξ_i^* се различни фактори на релаксација. За полесно пресметување, се воведуваат множители на Лагранж за да ги претворат гореспоменатите ограничени проблеми со оптимизација во двоен проблем.

Решението на равенката е следното:

$$f(x) = \sum_i^l (-\alpha_i + \alpha_i^*) K(x_i, x) + b, \quad (4.14)$$

каде, α_i и α_i^* се Lagrange множители што одговараат на векторите за поддршка (support vectors - SVs), l е број на SV, а $K(x_i, x)$ е јадрена (kernel) функција. Сепак кај SVR, различните јадрени функции имаат големо влијание врз резултатите.

Преферираната јадрена функција е радијално базираната функција (radial basis function - RBF):

$$K(x_i, x) = \exp(-\gamma \|x - x_i\|^2), \quad (4.15)$$

каде што γ е јадрен параметар. Во споредба со линеарното јадро, RBF јадрото проектира податочно множество во високо димензионално нелинеарно пространство, со што се овозможува ефикасно справување со нелинеарните врски помеѓу карактеристиките и ВР. Споредено со полиномното јадро, ова јадро има помалку параметри за подесување и ја намалува комплексноста на избраниот модел.

Со избор на соодветен фактор на казна ζ и јадрен параметар γ , функцијата на загуби ℓ може ефективно да се подобри. Сепак не постои генерално упатство за избор на параметрите. Брзото и ефикасно избирање на параметрите е клучот при предвидувањето на моделот.

4.3.5 Баесова класификација

Баесови класификатори се статистички класификатори кои се базираат на Баесовата теорема која ја предвидува веројатноста за одредена секвенца да припаѓа на одредена класа. Слично како дрвата на одлучување и селектираните класификатори во невронски мрежи, така и Баесовите класификатори (како наивен Баесов класификатор и Баесови мрежи), кога се применуваат на големи бази на податоци покажуваат висока точност и брзина.

Баесовата теорема ја дава постериорната веројатноста за одреден настан H , условен од X , $P(H|X)$. За тоа е потребна а priori веројатноста за настанот H , $P(H)$, постериорната веројатност за настанот X условен од настанот H , $P(X|H)$, и а priori веројатноста за настанот X , $P(X)$.

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)} \quad (4.16)$$

4.3.5.1 Наивен Баесов класификатор

За податочно множество каде што постои n - димензионален вектор, $X = (x_1, x_2, \dots, x_n)$, n атрибути, A_1, A_2, \dots, A_n и класа C како m - димензионален вектор,

$C = C_1, C_2, \dots, C_m$, наивниот Баесов класификатор функционира на следниов начин:

1. За која било n -точка, X , класификаторот ќе предвиди дека X припаѓа на класата која ја има највисоката постериорна веројатност, условена од X , ако и само ако $P(C_i | X) > P(C_j | X)$ за $1 \leq j \leq m, j \neq i$

Наивниот Баесов класификатор ќе превиди дека X припаѓа на класата C_i

На тој начин, $P(C_i | X)$ е максимизирана и според Баесовата теорема,

$$P(C_i | X) = \frac{P(X | C_i)P(C_i)}{P(X)}$$

2. Потребно е да се максимизира само $P(X | C_i)P(C_i)$ бидејќи $P(X)$ би била максимизирана.
3. Се прави претпоставка за условна независност на класите, намалувајќи ги пресметките во $P(X | C_i)$. За вредностите на атрибутите се претпоставува дека се условно независни еден од друг.
4. За предвидување на ознаката на класата, за секоја класа C_i се прави евалуација на $P(X | C_i)P(C_i)$. Целта е да, за n -точката X , се предвиди онаа класа C_i за која $P(X | C_i)P(C_i)$ има максимална вредност.

4.3.5.2 Баесова мрежа

Наивниот Баесов класификатор прави проценка на веројатноста, наместо стриктна класификација. За секоја вредност на класата, тој прави естимација на веројатноста за тоа дали дадена n -точка припаѓа на таа класа. Понатаму, за дадена класа, тој претпоставува дека атрибутите се условно независни еден од друг, на тој начин поедноставувајќи го пресметувањето.

Баесова мрежа, уште наречена и мрежа на верување е веројатносен графички модел кој претставува множество на случајни променливи и нивната условна независност со помош на насочен ацикличен граф (DAG). Баесовите мрежи се статистички базирана алтернатива која припаѓа на стохастички графички модели кои претставуваат множество од случајни променливи во јазлите и нивните условени зависимости преставени во врските помеѓу јазлите.

Овие мрежи ги комбинираат принципите од теорија на графика, теорија на веројатност, компјутерски науки и статистика. Ваквите мрежи ги специфицираат дистрибуциите на условната веројатност, кои овозможуваат условните независности меѓу класите да бидат дефинирани помеѓу подмножество од променливи.

Како и наивниот Баесов класификатор, Баесовите мрежи, исто така, користат статистички методи, со што нудат ефикасен и принципиелен пристап за избегнување на преклопувањето на податоци.

Ваквиот метод е развиен под претпоставка за номинални вредности (во случај на нумерички вредности, тие се предискретизираат) и отсуство на атрибути што недостасуваат. Баесовите мрежи имаат значително влијание врз моделирањето и анализата на податоците за пациентите. Главната предност е полесно толкување на резултатите и робустност за справување со податоците кои недостасуваат.

Баесовата мрежа може да ја претставува врската меѓу одредени болести и симптоми. Доколку се знаат симптомите, мрежата може да се искористи за да се пресмета веројатноста за присуство на некоја од различните болести. Постојат ефикасни алгоритми кои со помош на Баесова мрежа извлекуваат знаење за одредени проблеми (Vieira, 2011).

4.3.6 Асоцијативна класификација

Асоцијативна класификација е пристап кој ги користи техниките од асоцијативните правила за конструкција на класификациски системи.

Асоцијативните алгоритми бараат чести шаблони кои често се повторуваат и нивните соодветни асоцијации или корелациски правила со кои се карактеризираат релациите меѓу атрибутските услови и класните ознаки.

Во првиот чекор, алгоритмот бара чести *itemsets*, при што ги разгледува шаблоните од атрибут-вредност парови кои со повторување се појавуваат во податочното множество. Таму секој пар атрибут-вредност се смета за ставка-*item*, и на тој начин се формира зачестени *itemsets* (множества од ставки). Во вториот чекор, овие множества се анализираат за да се креираат асоцијациски правила.

4.3.6.1 RIPPER (JRip)

RIPPER (Repeated Incremental Pruning to Produce Error Reduction) прво е предложен од Cohen (Cohen, 1995) како оптимизирана верзија од IREP (Incremental Reduced Error Pruning) на (Fürnkranz & Widmer, 1994).

За да се извлечат правила на одлучување од податоченото множество, RIPPER врши поделба на тренинг податоците, создава единечно правило користејќи едно подмножество од податоците со употреба на евристички метод и потоа го поткаструва правилото користејќи едно подмножество од податоците. По создавање на множество од правила за класата, секое правило се преиспитува користејќи поткастрување за отстранување на грешки, и на тој начин се обезбедува точноста на секое правило, пред да се продолжи со генерирање на правилата за следната класа. Weka користи Java имплементација на RIPPER, позната како JRip. Главните предности на користењето на овој алгоритам се јасно множество на правила за класификација и неговата брзина на работа.

Во однос на класификацијата во повеќе истражувања правени се повеќе експерименти за да се одберат најсоодветните методи и од гледна точка на способноста за поделба, како и од времето поминато за да се донесе одлука. Според истражувањата естимацијата најчесто се фокусирана на следниве класификатори: J48, Random Forest, BayesNet и JRIP, кои се различни типови алгоритми за машинско учење и имаат претходни успешни примени во делот на кардиоваскурноста.

J48 алгоритмот претставува Weka имплементација на алгоритмот C 4.5, кој претставува најупотребуван класификатор на база на дрва развиен од страна на (Quinlan, 1993). Неговата имплементација се основа на пристап кој не се враќа назад. J48 користи два евристични критериуми за рангирање на можните тестови: добивка на информации (information gain) и предодредената рата на добивка (default gain ratio). По процесот на градење, секој тест на атрибутите, по должина на патеката (од коренот до листот) станува правило - предуслов, а класификацијата на јазол на листот станува правило последица. Дрвата на одлуки се лесни за толкување, способни да работат со вредности што недостасуваат,

како и со категорични и континуирани податоци, карактеристики што ги прават интересни и применливи во медицински цели.

Random Forest, BayesNet и JRIP веќе претходно се објаснети.

4.3.7 Симболичка регресија и генетско програмирање

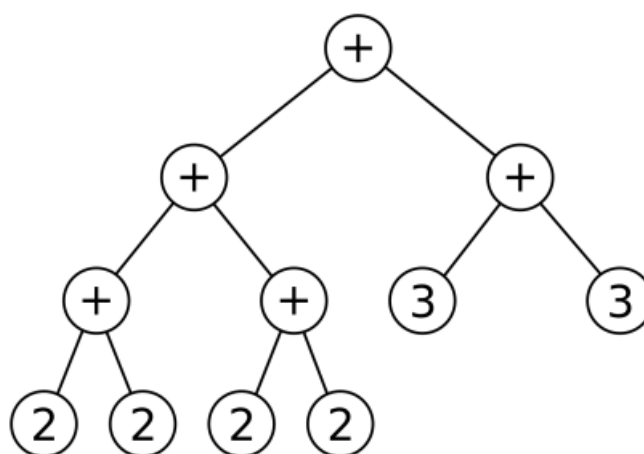
Симболичката регресија (Billard & Diday, 2002) е тип на регресивна анализа кој се обидува да креира математички изрази кои најдобро се вклопуваат во определено податочно множество. Ова е тесно поврзано со генетските алгоритми бидејќи ги користи истите концепти за развој на популација со користење на вкрстување и мутација. Идејата е прво да се креираат случајни математички изрази од претходно определено множество на градежни блокови, кои вклучуваат елементарни математички операции и нумерички операнди кои се нарекуваат терминали. Овие терминали се претставени со случајни константи или претходно добиени карактеристики од PPG-сигналот. Првичните случајни изрази ја претставуваат почетната популација и се експлицитни равенки за предвидување на систолниот (SBP) и дијастолниот притисок (DBP). Во следните итерации, изразите или поединците од населението се развиваат со користење на типични оператори за генетско програмирање, а тоа се селекција, вкрстување и мутација. Ова ги подобрува равенките за попрецизно предвидување на крвниот притисок. Изборот на најдобри поединци се заснова на фитнес-функцијата, која е избрана според средната квадратна грешка (Mean square error-MAE).

(G. Sannino et al., 2015) се првите кои го прифатиле ваквиот пристап за естимација на крвниот притисок. Нивниот предлог за ваков пристап, понатаму е изменет и проширен.

4.3.7.1. Поединци и популација

Целта на поединецот е да биде мост помеѓу контекстот на проблемот во реалниот живот и еволутивниот простор за решавање на проблемот. На тој начин, поединецот е дефиниран како математички израз или функција, создадена од претходно утврден збир на градежни блокови, што претставува експлицитна формула за пресметување на систолниот (SBP) или дијастолниот (DBP) притисок. Секој поединец може да биде претставена како структура на дрво, што се состои

од различен број и видови на градежни блокови, како што е прикажано на слика 4.9. Јазлите на дрвата одговараат на операторите, под дрвата на операндите, а листовите од дрвото одговараат на терминалите. Длабочината на дрвото е претходно ограничена, бидејќи не треба да се земаат предвид изрази над определена сложеност (Slapničar et al., 2018).



Слика 4.9: Изразот $((2 + 2) + (2 + 2)) + (3 + 3)$ претставен како дрво на изрази

Figure 4.9 The expression $((2 + 2) + (2 + 2)) + (3 + 3)$ represented as tree

Секоја карактеристика на PPG-сигналот заедно со множеството на основни математички функции и случајни константи, го сочинуваат множеството на градежните блокови за дрвата на изрази. Целото множество е дадено во табела 4.1. Некои оператори земаат еден влез (на пр. синус), додека други земаат два влеза (на пр. собирање). Тоа значи дека поединците не се строго бинарни дрва со изрази, бидејќи секој јазол може да има едно или две деца, во зависност од n -арноста на операторот што е претставен со јазолот. Целта на популацијата е да претставува различни можни решенија. Додека поединците се статични и не се менуваат, популацијата се менува, бидејќи е единица за еволуција. Почетната популација ја сочинуваат фиксни, релативно голем број случајно создадени поединци која потоа се развива низ генерации, со употреба на генетски оператори, биолошки инспирирани, кои се обидуваат да ја подобрат секоја последователна генерација според дадена избрана фитнес-функција.

Табела 4.1: Потенцијални градежни блокови на секој поединец (математички израз) во популацијата

Table 4.1 Potential building blocks of each individual (mathematical expression) in the population

Опис	Тип	Симбол	Арност (Бр на операции)
Константа	Оператор	C	0
Карактеристика	Оператор	x_i	0
Собирање	Оператор	$+$	2
Одземање	Оператор	$-$	2
Множење	Оператор	$*$	2
Делење	Оператор	$/$	2
Синус	Оператор	\sin	1
Косинус	Оператор	\cos	1
Природен логаритам	Оператор	\ln	1
Природен експонент	Оператор	e^{x_i}	1
Квадратен корен	Оператор	$\sqrt{x_i}$	1
Апсолутна вредност	Оператор	$ x $	1

4.3.7.2. Фитнес-функција

Улогата на фитнес-функцијата е да ја води еволуцијата, со претставување на побарувањата на кои мора да се прилагоди. Таа им доделува мерка на квалитет на поединците и на тој начин дефинира какво подобрување е потребно во контекст на еволуцијата. Во нашиот домен на проблемот, би сакале секој израз со PPG влезни променливи да врати точно предвидување за систолниот (SBP) и дијастолниот (DBP) притисок за тие вредности на PPG-сигналот. Затоа, средната апсолутна грешка (MAE) може да се користи како фитнес-функција, бидејќи добро

го опишува квалитетот на решението. Ни кажува колку добро предвидувањата се совпаѓаат со набљудуваните вредности и се дефинирани како:

$$MAE = \frac{\sum_{i=1}^n |y_i - \lambda(x_i)|}{N} = \frac{\sum_{i=1}^n e_i}{N} \quad (4.17)$$

каде y_i е набљудуваната вредност за крвниот притисок (BP), а $\lambda(x_i)$ е предвидената вредност за инстанцата x_i . Нивната разлика е грешката означена како e_i , а N е вкупниот број на предвидувања.

Сложеноста или големината на дрвото со изрази може да биде изводлива секундарна фитнес-функција, во ваков случај.

4.3.7.3 Механизам на селекција и генетски оператори

Генетските варијации претставуваат основа на еволуцијата. Улогата на генетските оператори е да се креираат нови поединци од оние кои се избрани од процесот на селекција. Споменатите механизми и оператори се следниве:

- **Селекција**

Селекцијата управува со процесот на репродукција, бидејќи ги избира специфичните поединци кои ќе бидат подложени на наведената репродукција. Селекцијата треба да фаворизира висококвалитетни поединци за да се создаде потенцијално супериорно потомство. Поединците со низок квалитет не смеат да бидат отфрлени, туку да им се даде шанса на низок избор. Ова се постигнува со употреба на механизам за турнирски избор, кој прво избира фиксен број поединци по случаен избор (со замена) и потоа ги избира најдобрите меѓу нив за репродукција, во однос на фитнес-функцијата. Дрвата со изрази, што одговараат на равенките со ниски MAE грешки, се очекува почесто да бидат избрани.

- **Вкрстување (Crossover)**

Вкрстувањето ги спојува информациите од два родители во два потомци. Поддрвото се избира случајно кај секое родителско дрво и

потоа овие поддрва се разменуваат. Длабочината на дрвата деца не смее да ја надминува претходно утврдената максимална длабочина. Ако дрвото дете е премногу длабоко, тоа е отфрлено и еден од родителите е избран и копиран во новата популација.

- **Мутација (Mutation)**

Мутацијата поттикнува генетската разновидност и секогаш претставува стохастички процес. Во дрвото дете се избира случаен внатрешен јазол и неговиот оператор се менува во друг случаен оператор со иста p -арност. Константите во листовите на дрвјата, исто така, можат да бидат мутирани.

- **Копирање (Copy)**

Мал процент од поединците родители е едноставно копиран во следната генерација. Понатаму, следи концептот на елитизам, што значи дека најдобрите поединци во популацијата секогаш се копираат без мутација. Ова осигурува дека новата популација ќе биде најмалку добра како и претходната, но се очекуваат и подобрувања.

Ваквите оператори се применуваат сè додека не се создаде нова генерација со ист број на поединци. Бидејќи е имплементиран елитизмот, се очекува приближување на кандидатските решенија кон добро решение (Giovanna Sannino et al., 2016).

Првичната генерација на 100 решенија (равенки), првиот пат се креира случајно. Решенијата се развиваат сè додека најдобриот поединец не го исполни потребниот праг за квалитет или додека 1 000 последователни генерации не покажат забележително подобрување во однос на фитнес-функцијата (MAE намалување поголемо од 1 mmHg). Големината на секое решение е ограничена на 100 градежни блокови. Кога процесот ќе запре, се разгледуваат најдобрите поединци создадени за време на еволуцијата.

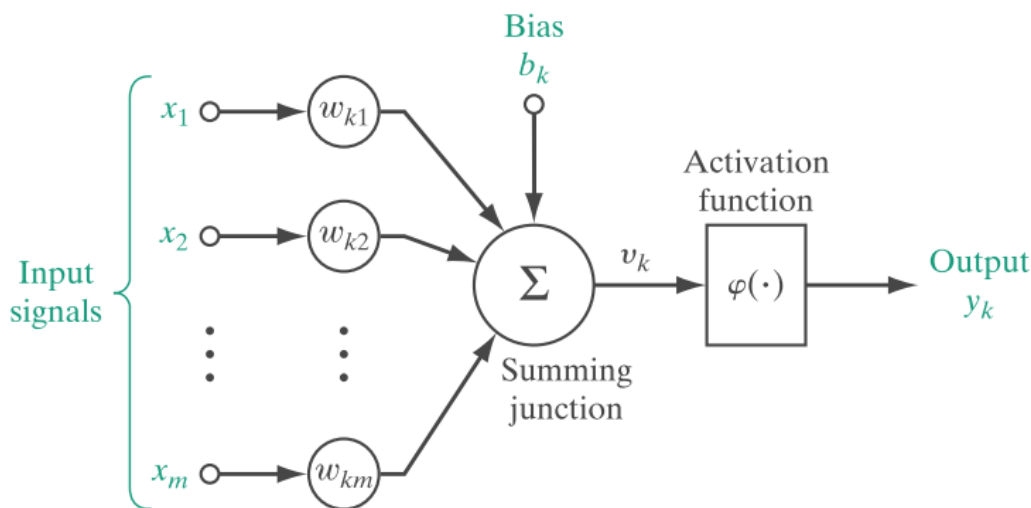
4.4 Невронски мрежи (Neural Network)

Вештачка невронска мрежа (Artificial Neural Network): За разлика од линеарната регресија, ANN може да се справи со случајот каде што карактеристичниот вектор

и целта имаат нелинеарна поврзаност. Естимацијата на крвниот притисок со ANN веќе е користена во повеќе истражувања како што се (Ludi Wang et al., 2018), (Rastegar et al., 2019). Перформансите на ANN многу зависат од бројот на скриени слоеви и скриени неврони.

Работата и истражувањата со NN произлегуваат од разбирањето на тоа како функционира човечкиот мозок при трансформирање на сложената група на влезови во едноставна акција или излез. На овој начин, исто како и човечкиот мозок, вештачката невронска мрежа претставува масовно паралелен дистрибуиран процесор, способен да чува големо експериментално знаење и да ги направи карактеристиките достапни за употреба.

Невронската мрежа се состои од неколку основни единици, наречени неврони. Модел на дигитален неврон заснован на моделот McCulloch-Pitts (Bishop, 2006) е претставен на слика 4.10, на која се претставени и главните компоненти на невронското дрво.



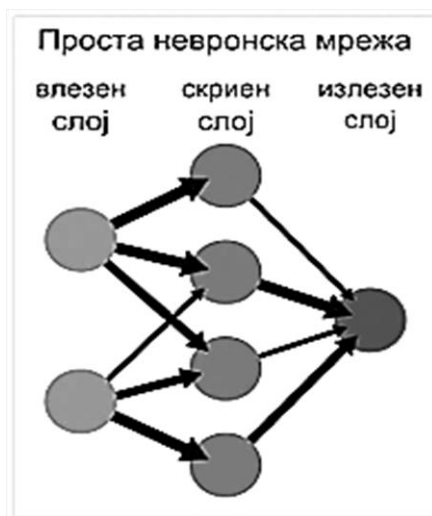
Слика 4.10: Модел на дигитален неврон заснован на моделот McCulloch-Pitts (Bishop, 2006)

Figure 4.10 Digital neuron model based on the McCulloch-Pitts model (Bishop, 2006)

Основните компоненти на невронското дрво се:

- Множество од синапси кои се карактеризираат со нивните тежини $w_{k1}, w_{k2}, \dots, w_{km}$, кои можат да се сметаат како значајност која ја има влезот врз целокупниот невронски излез.

- Функција за додавање, која ги додава сите влезни сигнали со нивните тежини според соодветните јачини на синапсите.
- Активациска функција за лимитирање на амплитудата од излезот на невронт.



Слика 4.11: Пример за едноставна невронска мрежа

Figure 4.11 Example of a simple neural network

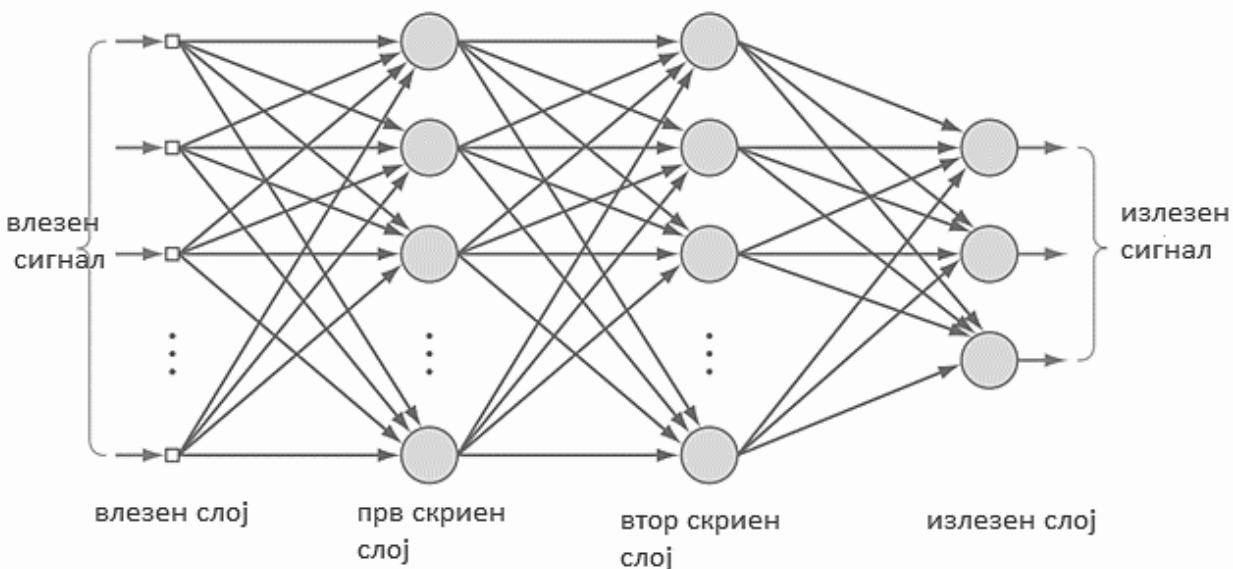
Вообичаено е да се вклучи влез на bias за да се зголеми или намали мрежниот влез на функцијата за активирање, ова помага да се контролира активациската функција, овозможувајќи и да се префрли во хоризонталната оска. Обично се претставува како фиксиран влез со вредност 1 со специфична тежина еднаква на bias. Невронот може да биде математички објаснет во следниве равенки:

$$v_k = \sum_{j=1}^m \omega_{kj} x_j + b_k \quad (4.18)$$

$$y_k = \phi(v_k) \quad (4.19)$$

Каде x_1, x_2, \dots, x_m се влезовите, $\omega_{k1}, \omega_{k2}, \dots, \omega_{km}$ се соодветните тежини, b_k е bias и ϕ е активациската функција.

Како што спомнавме и претходно, вештачката невронска мрежа е комбинација од едноставни невронски единици, кои обично се групираат во слоеви. На слика 4.12 е претставена повеќеслојна feed forward мрежа, каде што секој излез од некое ниво на невронт се зема како влез на невронт од следното ниво.



Слика 4.12: Пример за повеќеслојна feedforward мрежа

Figure 4.12 Example of a multilayer feedforward network

Кај ваквата мрежа можат да се идентификуваат три различни нивоа. Влезно ниво кое ги прима влезните податоци, еден или повеќе скриени слоеви каде се извршува главниот дел од пресметките и едно излезно ниво што го претвора резултат од скриениот слој во излез што може да биде класа (кога проблемот е класификација) или вредност (кога излезот е резултат на проблем со регресија). Меѓу другите, постојат и други типови мрежи како Radial Basis Functions NN, Recurrent NN и сл. Пред да може да се користат невронските мрежи за предвидување, неопходно е да се тренираат. Тренирањето генерално се изведува давајќи им на мрежата неколку примери на влезови за кои се познати излезите и се ажурираат тежините користејќи ги грешките. Најчесто користен алгоритам за итеративно да се изврши ваквото тренирање е алгоритмот на ширење наназад (back-propagation). Овој алгоритам може да се поедностави во неколку чекори:

1. Иницијализација на моделот: тежините се случајно иницијализираат;
2. Ширење на напред: влезните вектори се шират напред низ невронската мрежа;
3. Функција на загуба: излезните вредности се споредуваат со вистинските вредности, оваа споредба се прави со употреба на избрана функција на

загуба (на пр. Абсолютна разлика помеѓу пресметаните и вистинските вредности или сумата на квадратите на апсолутните грешки);

4. Диференцијација: пресметка на градиентот на функцијата на загуба, бидејќи тој ја дава брзината со која се менува функцијата.
5. Ширење на грешката на назад: Вредностите на грешката се шират назад низ мрежата за да се пресметаат вредностите на грешката на невроните во скриените слоеви. Градиентите на функција за загуба на скриените неврони можат да се решат со помош на правилото на синџирот;
6. Ажурирање на тежините: тежините се ажурираат со користење на правилото делта, кое се дефинира на следниот начин:

$$\Delta w_{ji} = \alpha(y_j^* - y_j)\phi'(h_j)x_i$$

7. Итерација: ваквата процедура се повторува додека тежините не се приближат.

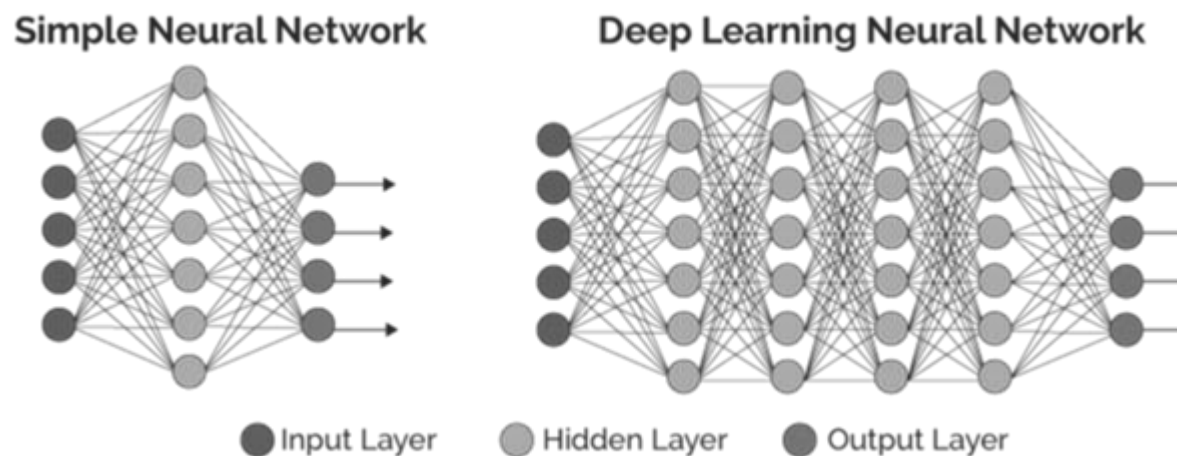
4.5 Длабоко учење (Deep learning)

Последните години, длабоко учење (deep learning) се појави како авангарден пристап кон машинското учење (machine learning). Пристапот не е нов бидејќи се основа на вештачки неврони мрежи (Artificial Neural Networks - ANNs) кои се присутни со децении.

ANN, како што веќе и претходно спомнавме се инспирирани од биолошките невронски мрежи и се состојат од меѓусебно поврзани едноставни елементи наречени јазли или неврони. Раната или почетната форма е повеќеслоен перцептрон (Multi Layer Perceptron - MLP), кој обично има мал број на неврони и сите врски се насочени од влезот кон излезот, што значи дека перцептрон е секогаш насочен напред. Невроните се елементи кои се процесираат кои содржат функција за активирање. Оваа активациска функција зема n влезни врски. Секоја влезна врска има соодветен параметар а тоа е тренинг тежина w_i . Функциите за активирање на невроните враќаат излезна врска, пресметана со функција за активирање. Најчести активациски функции се sigmoid, tangens hyperbolicus и rectifier (ReLU) (Goldberger et al., 2000).

Тренингот обично се извршува со користење на оптимизиран алгоритам за градиентно спуштање и обратно пропагирање. Тие се користат за да се прилагодат тежините на тренинг w_i врз основа на грешките кои се пресметуваат на излез. Грешките се дистрибуираат обратно низ нивоата во мрежата и тежините се прилагодуваат за да се намалат грешките во последователните итерации (повторувања).

Речењето на голем број на неврони во поврзана мрежа од слоеви ја гради вештачката невронска мрежа. Ако бројот на слоеви помеѓу влезниот и излезниот слој е голем (> 2), мрежата се смета за длабока (deep), како што е прикажано на слика 4.13.



Слика 4.13: Шематски приказ на „non-deep“ и deep ANNs

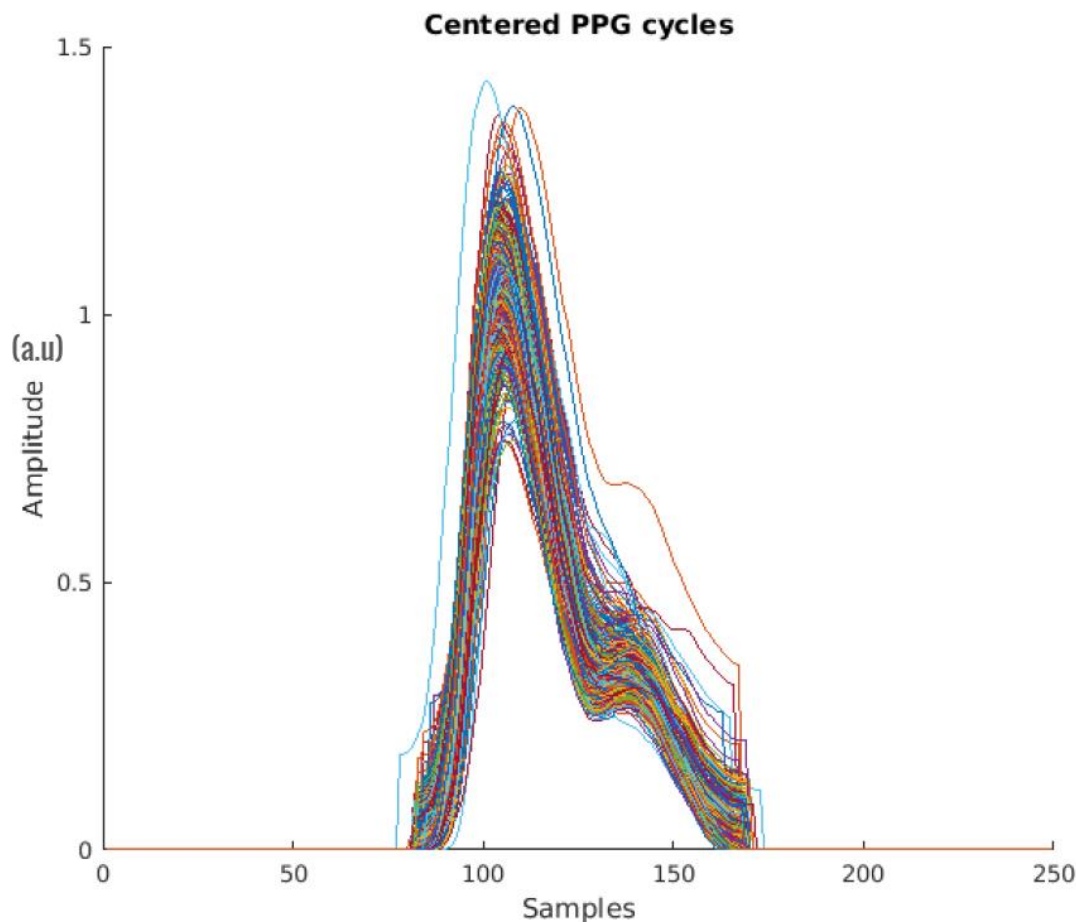
Figure 4.13 Schematic representation of „non-deep“ and deep ANNs

Со зголемување на обемот на податоците за обука и со зголемување на компјутерската процесорска моќ преку парарелизација, длабоките вештачки невронски мрежи го зголемуваат својот потенцијал. Екстремниот број на тежини кои требаа да се ажурираат за време на тренингот, може да претставува проблем, но со појавата на моќни графички процесорски единици (Graphics Processing Units - GPUs) кои овозможуваат голема паралелизација ова денес е лесно возможно. Денес модерните длабоки невронски мрежи може да имаат и до 1000 слоеви и милиони до милијарди неврони, што ги прави способни за моделирање на екстремно сложени нелинеарни врски.

Витален аспект по кој се разликуваат длабоките вештачки невронски мрежи од традиционалните методи на машинско учење е фактот дека тие се способни сами да донесуваат одлуки брз база на влезните необработени (raw) податоци. Во минатото, креирањето на карактеристики вообичаено беше во доменот на истражувачите кој потоа ги внесуваа овие карактеристики во алгоритмот за учење. Длабоките невронски мрежи им овозможуваат на истражувачите едноставно и единствено да ги снабдуваат мрежите со голема количина на обележани (етикетирани) влезни податоци, а потоа тие самостојно може да изведат карактеристики и да учат самостојно. Покрај ова рачно воведените карактеристики може да ги унапредат перформансите на моделот. Исто така ваквите мрежи може да комбинираат необработен сигнали рачно поставени карактеристики, што ги прави уште помоќни (Goodfellow et al., 2016).

Друга предност е фактот дека длабоките невронски мрежи се способни за моделирање на екстремно сложени нелинеарни врски, со што имаат поголема експресивна моќ во споредба со некои традиционални методи.

Во (Slapničar, 2018) се разгледува можноста за индивидуално користење на чист необработен PPG сигнал и рачно дефинирани карактеристики како и нивна комбинација. Барањата кај необработените е тоа да секоја истанца има иста должина. Ова е лесно да се осигури кога се користат карактеристики, но е потешко при користење на необработени PPG сигнали. Ова, тие го решаваат со центрирање на секој PPG циклус во прозорец од 2 секунди, што е доволно долго за секогаш да се содржи еден пулс, слика 4.14.



Слика 4.13 500 PPG циклуси каде секој е центриран во прозорец од 2 секунди
(Slapničar et al., 2018)

Figure 4.13 500 PPG cycles, each is centered in a 2 second window (Slapničar et al., 2018)

И покрај споменатите предности на длабоките невронски мрежи, тие не можат многу лесно да се користат и моделираат надвор од некој определени рамки (out of the box). Покрај големата процесиравка комплексност, голем предизвик лежи и во определување на оптималните параметри за мрежата да може ефикасно да учи. Топологијата на невронската мрежа е непозната, бидејќи таа може да биде длабока (deep) или плитка (shallow), широка (wide) или тесна (narrow). Генерално се зема дека ако е подлабока тогаш е подобро. Сепак, ова не е секогаш точно и не е задолжително да се зема како факт. Дополнително, колку е подлабоко толку се зголемува комплексноста. Следната непозната е видот на активациската функцијата која треба да се користи како и типот на алгоритмот за оптимизација

кој треба да осигури конвергирање кон локалниот минимум. Типовите на нивоата се следниот параметар од интерес. Бидејќи постојат различни типови на слоеви (целосно поврзани, конволуциски, рекурентни, Long-Short-Term-Memory итн.) погодни за различни проблеми. На крај треба да се изберат голем број на други параметри, како стапката на учење, потенцијални параметри за маневрирање и регулирање, со помош на кои се осигурува дека мрежата нема да отстапи од нејзината намена.

Влезното ниво на мрежата може да добие пресметани карактеристики од централните необработени висококвалитетни PPG сигнали или може да се прави комбинација. Излезниот слој на мрежата се состои од два неврона, едниот за SBP, а другиот за DBP. Некои дополнителни параметри кои се избираат во зависност од топологијата ја определуваат стапката на учење и типот на слоевите.

4.6 Евалуација

Главна карактеристика на моделот на машинско учење е колку добро може да генерализира нови невидени податоци. Кога трениран модел не може да обезбеди најдобри резултати при оценување, можно е моделот да е преполнет со податоци во тренинг множеството.

За оценување на ова сценарио, множеството податоци вообичаено се дели на две или три групи.

- Тренинг множество: податоците кои ќе бидат искористени за тестирање на моделот од машинско учење;
- Множество за потврдување: едно „дополнително“ множество кое може да се искористи за подесување на хиперпараметрите (параметри поставени пред процесот на учење) со нови невидени податоци;
- Тест множество: множество од невидени примери кои се користат само за проценка на перформансите на моделот.

Еден важен фактор при оценување на моделот е тест множеството од податоци да биде избрано репрезентативно за множеството податоци во целина и доволно големо за да се добие статистички значајни резултати.

Неколку метрики можат да се користат за оценување на моделот, бидејќи оваа работа се заснова на регресиски модели, вообичаените метрики кои се користат за оценување на моделот се:

Кога n е бројот на примероци, y е реалната вредност, а y^* е проценетата вредност:

- Средна грешка (ME)

$$ME = \frac{1}{n} \sum_{i=1}^n y - y^* \quad (4.20)$$

- Средна апсолутна грешка (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - y^*| \quad (4.21)$$

- Средно квадратна грешка (RMSE): осетлив на големи грешки

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - y^*)^2} \quad (4.22)$$

- Коефициент на детерминација (R^2): обезбедува индикација колку се совпаѓаат предвидените вредности од конкретните вредности

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (4.23)$$

каде $SS_{tot} = \sum_i (y_i - \bar{y})^2$, $SS_{res} = \sum_i (y_i - y_i^*)^2$, y_i е измерената вредност, y_i^* е

проценетата вредност и \bar{y} е средната вредност на y .

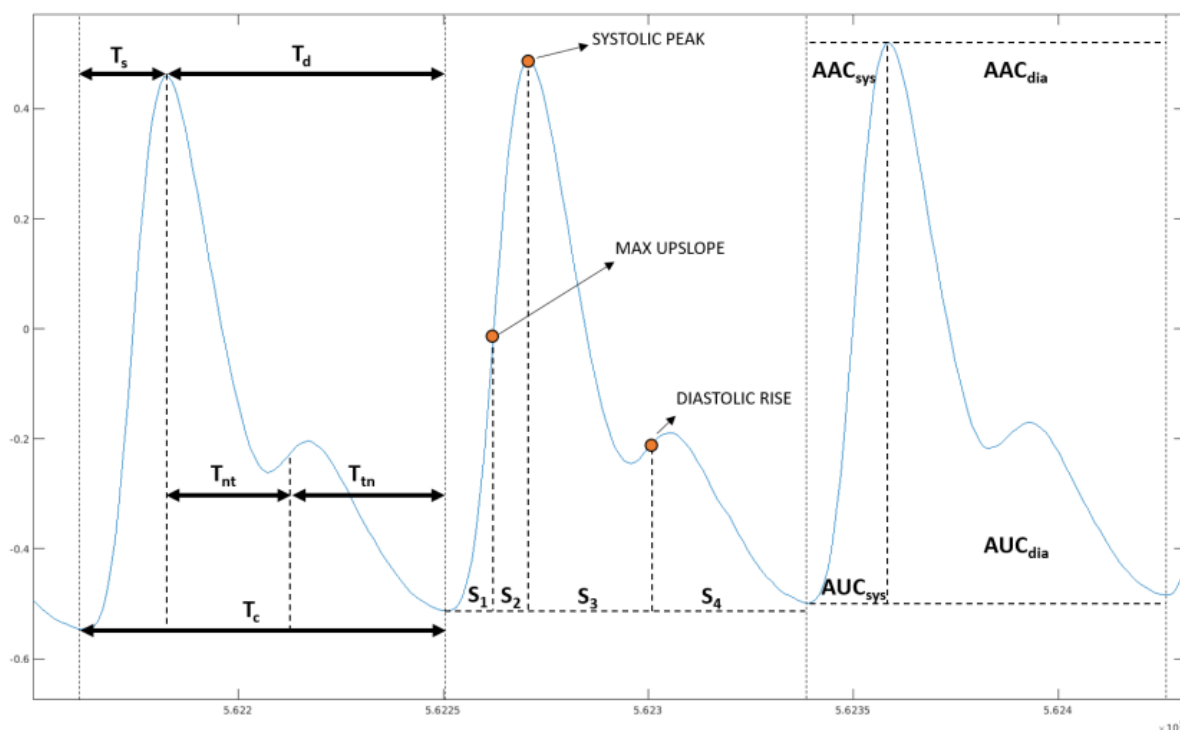
5. Основни карактеристики што го опишуваат PPG-сигналот

PPG карактеристиките најчесто ја опишуваат морфологијата на еден индивидуален циклус во временски домен, користејќи траење на одредени карактеристични форми во самиот PPG бран, или површни под или над одредени делови од бранот. Од PPG-сигналот можат да извлечат голем број на карактеристики од временскиот домен, кои ги вклучуваат споменатите траења и површини од бранот. Според (Slapničar et al., 2019) (Slapničar et al., 2018) овие карактеристики се дадени на табела 5.1., и соодветно прикажани на слика 5.1.

Табела 5.1: Карактеристики на PPG-бранот во временски домен користени кај

Table 5.1 PPG wave features in time domain used in (Slapničar et al., 2018)

Карактеристика	Опис
T_c	Траење на циклусот
T_s	Време од почетокот на циклусот до систолниот врв
T_d	Време од систолниот врв до крајот на циклусот
T_{nt}	Време од систолниот врв до дијастолниот раст
T_{tn}	Време од дијастолниот раст до крајот на циклусот
S_1	Површина под кривата (AUC) од почетокот на циклусот до максималната upslope точка (точка – временска средина од почетокот на циклусот до систолниот врв)
S_2	Површина под кривата (AUC) од максималната upslope точка до систолниот врв
S_3	Површина под кривата (AUC) од систолниот врв до дијастолниот раст
S_4	Површина под кривата (AUC) од дијастолниот раст до крајот на циклусот
AUC_{SYS}	$S_1 + S_2$
AAC_{SYS}	Површина над кривата (AAC) од почетокот на циклусот до систолниот врв
AUC_{DIA}	$S_3 + S_4$
AAC_{DIA}	Површина над кривата (AAC) од систолниот врв до крајот на циклусот



Слика 5.1: Приказ на карактеристиките на PPG-бранот во временски домен (Slapničar et al., 2018)

Figure 5.1 PPG wave features in time domain (Slapničar et al., 2018)

Како дополнување на карактеристиките кои ја опишуваат морфологијата на PPG-бранот, се издвоени и следните карактеристики кои ја опишуваат состојбата на кардиоваскуларниот систем.

1. AI - Augmentation Index (Индекс на зголемување): е мерка за одбивање на бранот од артериите.

$$AI = \frac{\text{diastolic rise amplitude}}{\text{systolic peak amplitude}} \quad (5.1)$$

2. LASI - Large Artery Stiffness Index е мерка, односно индикатор за артериската крутост којашто е означена како T_{nt} , на слика 5.1. и во табела 5.1.

Со користење на комплексна анализа и фреквенциски домен, произлегуваат нови множества на карактеристики.

1. Карактеристиките од комплексната анализа: signal complexity и mobility се пресметуваат за PPG-сегмент од 30 секунди, на следниот начин.

$$E_0 = \sqrt{\frac{\sum_{i=1}^N PPG(i)^2}{N}} \quad (5.2)$$

$$E_1 = \sqrt{\frac{\sum_{j=2}^{N-1} d_j^2}{N-1}} \quad (5.3)$$

$$E_2 = \sqrt{\frac{\sum_{k=3}^{N-2} g_k^2}{N-2}}, \quad (5.4)$$

каде x е PPG-сигналот, N е вкупниот број на примероци во дадена временска низа, d е прв извод од x , а g е втор извод од x .

$$Complexity = \sqrt{\frac{E_2^2}{E_1^2} - \frac{E_1^2}{E_0^2}} \quad (5.5)$$

$$Mobility = \frac{E_1}{E_0} \quad (5.6)$$

2. Карактеристики од фреквентниот домен: Амплитуди и фази од претставувањето на PPG сигнал од 30 секунди во фреквентен домен, кои се дадени како:

$$PPG_{FREQ} = \text{fft}(PPG_{TIME}), \quad (5.7)$$

Амплитудите се добиваат како:

$$|PPG_{FREQ}[k]| = \sqrt{\left(\text{Re}(PPG_{FREQ}[k])\right)^2 + \left(\text{Im}(PPG_{FREQ}[k])\right)^2}, \quad (5.8)$$

а фазите се пресметуваат како:

$$\angle PPG_{FREQ}[k] = \arctan\left(\frac{\text{Im}(PPG_{FREQ}[k])}{\text{Re}(PPG_{FREQ}[k])}\right), \quad (5.9)$$

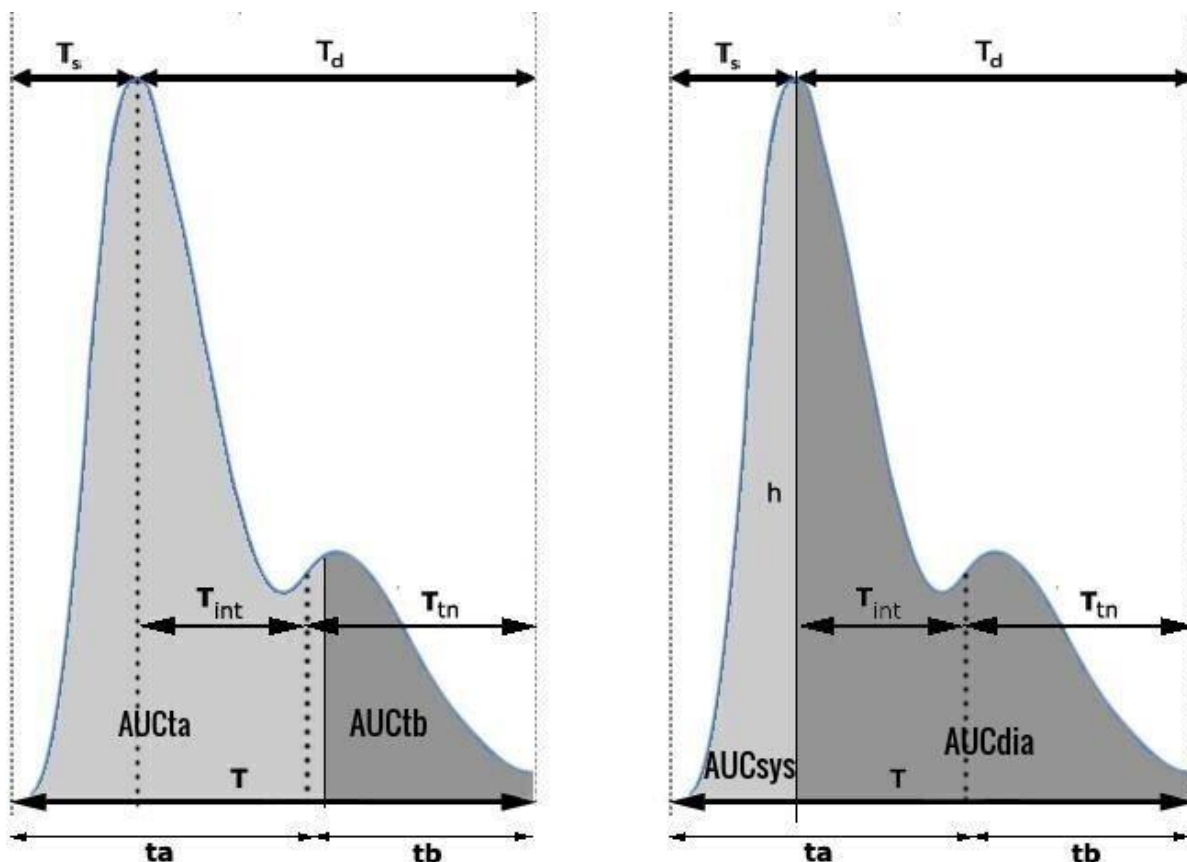
Каде што $\text{Re}(PPG_{FREQ}[k])$ е реалниот, а $\text{Im}(PPG_{FREQ}[k])$ е имагинарниот дел од фреквентните компоненти на PPG-сигналот, откако е направена брза Фуриева трансформација.

Дел од овие карактеристики се користени во (Slapničar et al., 2018). Ние дополнително додаваме други морфолошки карактеристики кои се полесни за извлекување од изгледот на самиот бран, а тоа се времињата t_a и t_b , каде што t_b се поклопува со T_{tn} и висината h . Дополнително ги имаме површините под соодветните криви, AUC_{ta} и AUC_{tb} , преставени на слика 5.2.

Табела 5.2: Карактеристики на PPG-бранот во временски домен користен и во нашиот случај

Table 5.2. PPG wave features in time domain, used in our case

Карактеристика	Опис
T	Траење на циклусот
T_s	Време од почетокот на циклусот до систолниот врв
T_d	Време од систолниот врв до крајот на циклусот
T_{nt}	Време од систолниот врв до дијастолниот раст односно до секундарниот врв (dicrotic peak)
T_{tn}	Време од дијастолниот раст (dicrotic peak) до крајот на циклусот
T_a	Време од почетокот на циклусот до dicrotic peak
T_b	Време од dicrotic peak до крајот на циклусот
AUC_{ta}	Површина под кривата (AUC) почетокот на циклусот до dicrotic peak
AUC_{tb}	Површина под кривата (AUC) dicrotic peak до крајот на циклусот
AUC_{sys}	Површина под кривата (AUC) од почетокот на циклусот до систолниот врв
AUC_{dia}	Површина под кривата (AUC) од систолниот врв до крајот на циклусот
h	Висината од дијастолната вдлабнатина до систолниот врв



Слика 5.2: Морфолошки карактеристики на PPG-сигналот користени во нашиот случај

Figure 5.2 Features of PPG wave used in our case

Значи, сите морфолошки карактеристики кои ние ги користиме, соодветно се преставени во табела 5.2 и соодветно на слика 5.2.

5.1 Анотацијана PPG сигналите

За да ги издвоиме потребните карактеристики кои ги спомнавме претходно, прво треба да ги најдеме точките кои ги претставуваат соодветно систолниот врв, дијастолната вдлабнатина, секундарниот врв (dicrotic peak) и секундарната вдлабнатина (dicrotic notch). За таа цел го користиме алгоритамот AMPD (автоматскиот, со повеќе нивоа базиран алгоритам за детекција на врвови (automatic multiscale-based peak detection-AMPD) (Scholkmann et al., 2012). И дополнителен алгоритам кој се користи за анотација на крвниот притисок имплементиран во Matlab од страна на Alexandre Laurin (Laurin, A. 2017).

5.1.1 Алгоритми за детекција на врвови

5.1.1.1 AMPD алгоритам

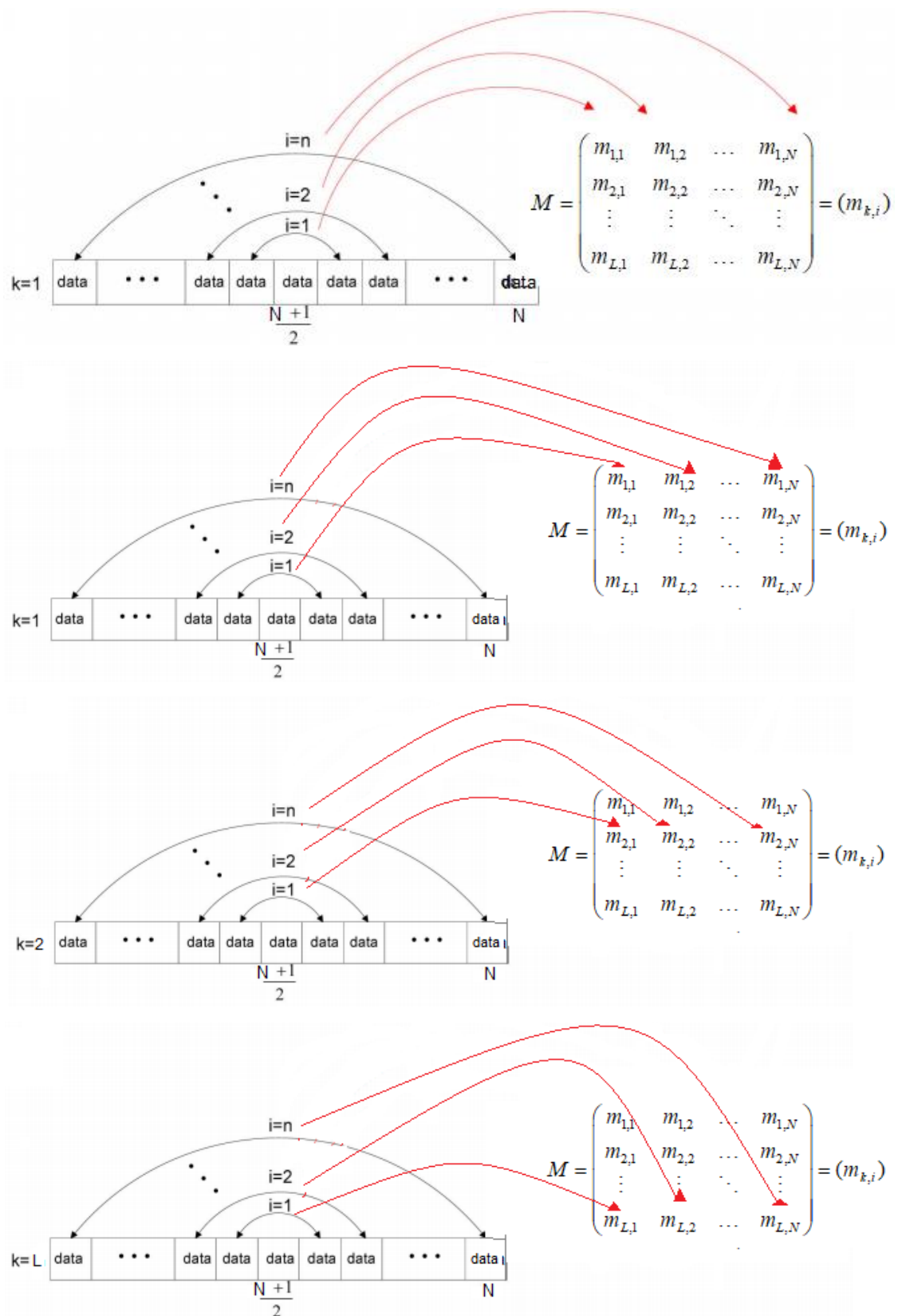
Треба да е даден $x = [x_1, x_2, \dots, x_i, \dots, x_N]$ како униваријантен рамномерно семплиран сигнал, кој содржи компонента на периодични или квазипериодични врвови. Првиот чекор на автоматскиот multiscale- базиран алгоритам за детекција на врвови (automatic multiscale-based peak detection-AMPD) се состои од пресметување на скалограм за локални максимуми (local maxima scalogram-LMS). За таа цел, сигналот x е одреден прволинеарно, т.е., се пресметуваат најмалите квадрати кои се вклопуваат на правата линија и се одземаат од x . Потоа, локалниот максимум на сигналот x се определува со користење на пристапот на движечки прозорец, при што се менува должината на прозорецот w_k , $\{w_k = 2k \mid k=1, 2, \dots, L\}$, каде k е k -тата скала на сигналот $L = \lceil N/2 \rceil - 1$ каде $\lceil z \rceil$ е функција која го дава најмалиот цел број не помал од z . Ова се реализира за секоја скала k и за $i = k + 2, \dots, N - k + 1$ според равенството:

$$m_{k,i} = \begin{cases} 0, & \text{ако } x_{i-1} > x_{i-k-1} \wedge x_{i-1} > x_{i+k-1} \\ r + \alpha, & \text{инаку} \end{cases} \quad (5.10)$$

Каде r е рамномерно распределен случаен број во опсегот $[0, 1]$, а α е константен фактор ($\alpha = 1$). За $i = 1, \dots, k + 1$ и за $i = N - k + 2, \dots, N$, вредноста $r + \alpha$ се доделува на $m_{k,i}$. Ваквите операции од претходното равенство резултираат во матрица

$$M = \begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,N} \\ m_{2,1} & m_{2,2} & \dots & m_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{L,1} & m_{L,2} & \dots & m_{L,N} \end{pmatrix} = (m_{k,i}) \quad (5.11)$$

Каде што k -тата редица ја содржи вредноста на прозорецот со должина w_k . На тој начин, сите елементи од матрицата M ($L \times N$) се наоѓаат во опсегот $[0, 1 + \alpha]$. Матрицата M се нарекува LMS на сигналот x .



Слика 5.3: Постапка за добивање на матрица M (LMS на сигналот x)

Figure 5.3 Procedure for obtaining matrix M (LMS of the signal x)

Вториот чекор од алгоритмот опфаќа собирање по редници на LMS матрицата M

$$\gamma_k = \sum_{i=1}^N m_{k,i}, \text{ за } k \in \{1, 2, \dots, L\} \quad (5.12)$$

Векторот $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_i, \dots, \gamma_L]$ содржи информации за дистрибуцијата на нулите која е зависна од слоевите (скалите) (а со тоа и информација за локалните максимуми). Глобалниот минимум на γ , $\lambda = \arg \min(\gamma_k)$, го претставува слојот со најголемиот локален максимум. Оваа вредност λ се користи во третиот чекор на алгоритмот за да ја измени LMS матрицата M со отстранување на елементите $m_{k,i}$, за кои $k > \lambda$, преку што се добива нова матрица $\lambda \times N$, $M_r = (m_{k,i})$, за вредности на $i \in \{1, 2, \dots, N\}$ и $k \in \{1, 2, \dots, \lambda\}$.

Во последниот чекор на алгоритмот, се наоѓаат врвовите со пресметување на стандардна девијација на матрицата M_r по колони, според равенството:

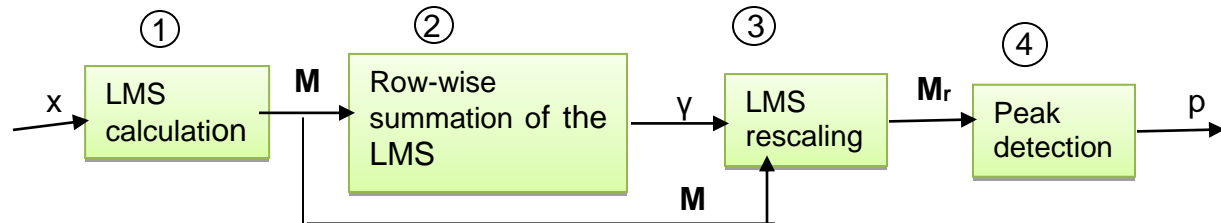
$$\sigma_i = \frac{1}{\lambda - 1} \sum_{k=1}^{\lambda} \left[\left(m_{k,i} - \frac{1}{\lambda} \sum_{k=1}^{\lambda} m_{k,i} \right)^2 \right]^{\frac{1}{2}} \text{ за } i \in \{1, 2, \dots, N\} \quad (5.13)$$

и се наоѓаат сите индекси i за кои важи $\sigma_i = 0$. Овие вредности се сместуваат во векторот $p = [p_1, p_2, \dots, p_q, \dots, p_{\hat{N}}]$, каде \hat{N} е вкупниот број на врвови во сигналот x . p се однесува на индексите на детектираните врвови. Емпириски е утврдено дека за влезни квазипериодични сигнали, највисоката фреквенција (f_{\max}) на осцилациите не треба да биде поголема од четири пати од најниската фреквенција (f_{\min}), односно $f_{\max} < 4f_{\min}$.

На ваков начин, користејќи го AMPD алгоритамскиот шаблон, секоја вредност од дадениот сигнал $x = [x_1, x_2, \dots, x_i, \dots, x_n]$ се детектира како врв во случај

кога стандардната девијација по колони на матрицата M_r , го задоволува критериумот $\sigma_i = 0$.

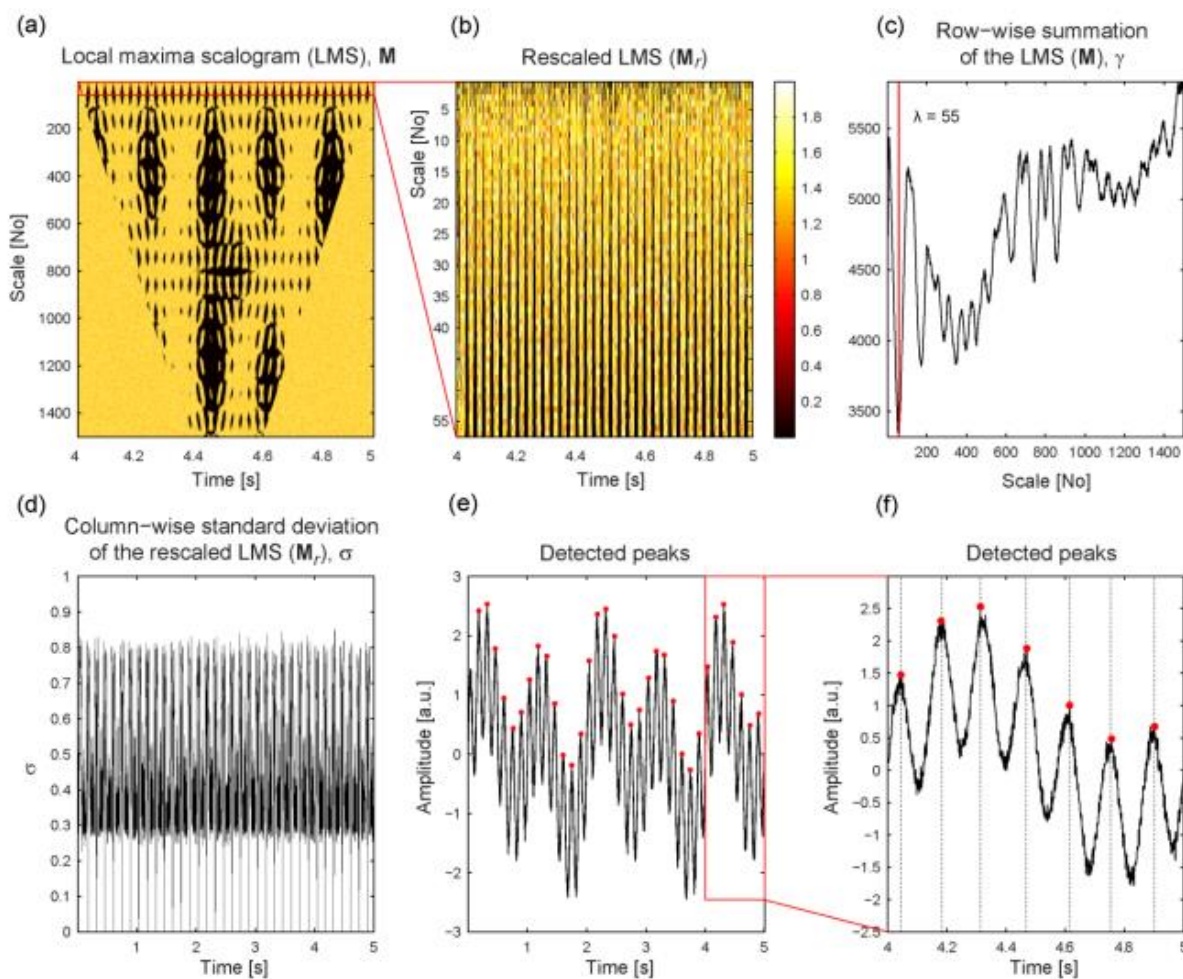
На сликата 5.4, визуелно се претставени чекорите на AMPD-алгоритмот, а на сликата 5.5 е претставен пример на примена на AMPD-алгоритмот на симулиран повеќекомпонентен сигнал со шум, x , дефиниран на следниов начин $x = [x_1, x_2, \dots, x_i, \dots, x_n]$, каде $x_i = \{a \sin(2\pi f_1 / f_s) + b \sin(2\pi f_2 / f_s) + c \sin(2\pi f_3 / f_s) + d \varepsilon_i\}$, за $i \in \{0, 0.1, \dots, N\}$, каде $f_s = 800\text{Hz}$, $f_1 = 10\text{Hz}$, $f_2 = 70\text{Hz}$, $f_3 = 5\text{Hz}$, $a = 1$, $b = 1$, $c = 0.5$, $d = 0.1$ и ε е нормално распределен случаен број во опсегот $[0, 1]$.



Слика 5.4: Тек на AMPD-алгоритмот

Figure 5.4 AMPD algorithm flow

Ние го користиме овој опишан алгоритам, а како помошен го користиме и алгоритмот (Laurin, 2017), за изнаоѓање на дикротик врвот. Со користење на двата алгоритми, ние ги добиваме точките кои ни се од интерес. Тие точки, понатаму, ни помагаат да добиеме некои од карактеристиките кои се претходно споменати.



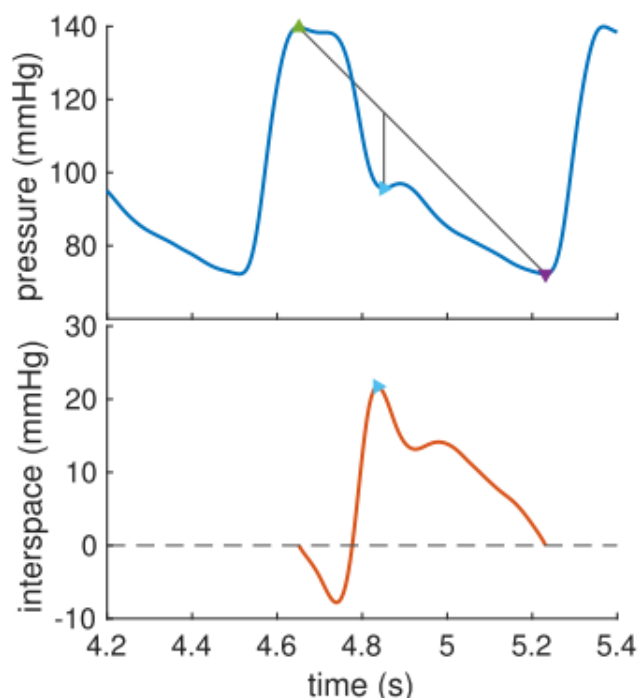
Слика 5.5: Пример од употреба на AMPD-алгоритмот врз симулиран сигнал (Scholkmann et al., 2012)

Figure 5.5 Example of using the AMPD algorithm on a simulated signal (Scholkmann et al., 2012)

5.1.1.2 Алгоритам за анотација на BP

Со помош на овој алгоритам (Laugin, 2017), се определуваат примарните минимуми (foot на бранот) и примарните максимуми (систолните врвови), секундарните локални минимуми (dicrotic notch) и секундарните врвови (dicrotic peaks). Техниката која се користи во алгоритмот е опишана во (Pan & Tompkins, 1985). Истовремено се користат критериуми кои се опишани во (Sun et al., 2006). Алгоритмот е наменет за сигнал за крвен притисок, но ние во нашиот случај го применуваме за PPG-сигналот, кој има сличен изглед на бранот. Значи, како влез користиме континуиран PPG-сигнал во временски домен и дополнително ја

наведуваме фреквенцијата на семплирање. Како излез ги добиваме точките за минимумите и за максимумите кои претходно ги споменавме (foot Index, systolic Index, notch Index и dicrotic Index). Според алгоритмот, foot индексот е дефиниран како точка каде вториот извод од временските серии е највисок во секој интервал каде подвижниот просек од вториот извод е поголем од прилагодливиот праг. Систолниот врв е дефиниран како максимум на бранот по определениот foot индекс, во прозорец со радиус $1/8$ секунда. Dicrotic notch е дефиниран како минимум од одземањето на сигналот и права линија која ги поврзува индексот на систолниот врв и дијастолниот индекс. Прикажано на слика 5.6.



Слика 5.6: Определување на dicrotic notch според (Le Gall et al., 2017)

Figure 5.6 Determination of dicrotic notch according to (Le Gall et al., 2017)

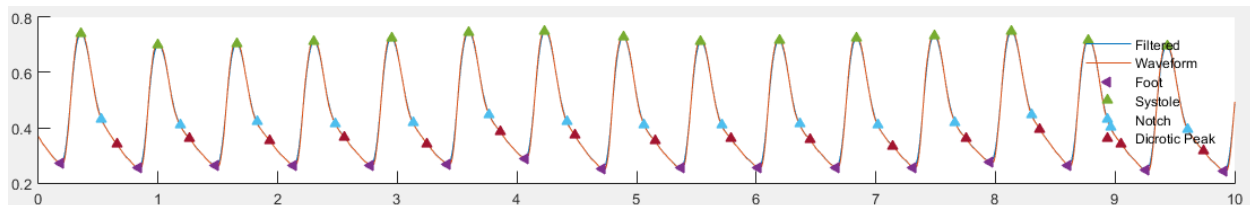
За t_i^s и t_{i+1}^f времето на систолниот врв и подножјето на 2 последователни отчукувања, времето на dicrotic notch t_i^d е дефинирано како време t што ја максимизира функцијата

$$d(t) = \max \left(S \cdot t + \frac{P(t_i^s)}{S \cdot t_i^s} - P(t) \right), \quad (5.14)$$

каде што S е наклонот $\frac{P(t_i^s) - P(t_{i+1}^f)}{t_i^s - t_{i+1}^f}$, P е притисок и $t \in [t_i^s, t_{i+1}^f]$.

Индексот на Dicrotic врвот се дефинира како минимум од вториот извод на временската серија која следи по dicrotic notch, во однос на прозорец со радиус RR/5 секунди. (RR е среден интервал на отчукување на срцето, кој се пресметува со foot индексите).

Со користење на овие алгоритми се добиваат определени точки кои понатаму служат да се извлечат потребните карактеристики. Пример за изглед на резултатот од алгоритмот е прикажан на слика 5.7.



Слика 5.7: Приказ на резултатот со користење на алгоритмот (Laurin, 2017)

Figure 5.7 Presentation of the result using the algorithm (Laurin, 2017)

Но, сигналите секогаш не се така чисти за да може јасно да се определуваат индексите на dicrotic notch и dicrotic peak, во такви случаи полесно е самоопределувањето само на врвовите и минимумите. За таа цел користиме комбинација од двата претходно споменати алгоритми.

5.1.2 Определување на карактеристиките и анотација на сигналот

Откако се најдени соодветните точки, понатаму ги определуваме потребните карактеристики дефинирани во табела 5.2. Времето T го определуваме како растојание помеѓу два foot индекса. Времето T_s , како растојание од foot индекс до систолниот индекс. Времето T_d , како растојание од систолниот индекс до следен foot индекс. Времето T_{nt} , како растојание од систолниот индекс до следен

индексна dicrotic peak. Времето T_{tn} , растојание од индексот на dicrotic peak доследниот foot индекс. Времето t_a , како растојание од foot индекс до dicrotic peak и t_b исто како T_{tn} .

Следно, се наоѓаат површините под кривите на самиот сигнал. AUC_{ta} карактеристиката се пресметува како површина под кривата што ја зафаќа бранот на PPG, од foot индекс до dicrotic peak. AUC_{tb} , како површина под кривата што ја зафаќа бранот на PPG-сигналот од индексот на dicrotic peak до следниот foot индекс. Карактеристиката AUC_{sys} се пресметува како површина што ја зафаќа кривата од PPG-сигналот од foot индексот до систолниот индекс и AUC_{dia} , како површина под истата крива само од точка на систолниот индекс до следен foot индекс. Карактеристиката за висина, односно параметарот h , се добива како висинско растојание од foot индекс, до следниот систолен индекс.

Откако се извлечени карактеристиките од PPG-сигналот, следно прашање е колкав дел од сигналот треба да се користи како податок кој подоцна ќе го искористиме за машинското учење.

Ако земеме сигнал во траење од 1 минута, и ги извлечеме карактеристиките од него, потоа тие карактеристики ги уредиме за да добиеме просечни вредности како карактеристики за тој сигнал, тогаш би можеле да добиеме поголема прецизност, но само ако сигналот е целосно чист и без прекини. Исто така, за негова обработка би ни било потребно и повеќе време. Но, во целата минута сигналот не е секогаш без прекини чист, бидејќи PPG-сигналот може брзо да се менува и многу е подложен на шум. Во таков случај може да се добие ситуација во која ќе добиеме вредности за карактеристиките кои не се реални. Ако се земе само една периода од бранот, во тој случај брзо ќе се обработи, но таа една периода недоволно го отсликува целиот бран бидејќи може да биде исклучок од останатиот дел. За да се прилагодиме на двете крајности ние се одлучивме да земаме 5 последователни периода, односно тоа би значело по 5 од сите претходно спомнати карактеристики. На тој начин, може да се даде доволна точна претстава за бранот, а од друга страна времето за пресметка би било релативно брзо. Од извлечените карактеристики, бидејќи од сите имаме по 5, ние наоѓаме

просечни вредности и тоа би било еден податок како влез во некои од формите за машинско учење.

Но, тука се поставува друго прашање, кои 5 последователни периоди на бранот ќе ги земеме предвид. Ако се првите 5, тогаш можеби тие не се целосни или пак, повторно не го отсликуваат добро целиот бран. Целта е тие 5 последователни периоди да бидат дел од бранот кој е чист и без шум. Само на таков начин, лесно од него би можело да се извлечат потребните карактеристики. Тие особини, не значи дека со сигурност ги имаат првите 5 периоди од бранот. Затоа ние се одлучивме од бран во траење, од максимум 1 минута, да ги определиме сите периоди и да ги земеме предвид оние 5 последователни периоди кои имаат најмали варијации меѓу себе. Но, тоа не е доволен критериум, бидејќи периодите можат да им бидат приближно исти, но сличноста на бранот во тие периоди да биде сосема мала. Затоа, како дополнителен критериум ја земаме сличноста на бранот во 5 последователни приближно исти временски периоди, што ја определуваме преку коефициент на корелација помеѓу секои два периода од бранот. Затоа, како прв чекор издвојуваме повеќе групи од 5 приближно исти периоди, кај кои разликата во периодите е најмала, а потоа кај сите од нив со помош на коефициент на корелација ја определуваме сличноста. На крај, ја извлекуваме онаа група со најголеми коефициенти на корелација.

5.2 Влезно податочно множество

Откако се извлечени сите потребни карактеристики од PPG-сигналот кои можат да послужат како податок за учење во некои од методите на машинското учење, следен чекор беше дефинирањето на влезно податочно множество со доволен број на податоци. Освен PPG-сигналот, во процесот на учење кај методите за машинско учење ни се потребни и соодветните вредности за систолниот и за дијастолниот крвен притисок, кој временски одговара со 5-те погодни периоди од кои ги извлекуваме потребните карактеристики. Потребните податоци ги земаме од [Physionet.org](https://physionet.org).

PhysioNet е извор на ресурси за физиолошки сигнали. На овој сајт се достапни голем број на клинички бази на податоци. Ние работевме со MIMIC-III

базата на податоци. Оваа база се состои од MIMIC-III Waveform база на податоци, MIMIC-III клиничка база на податоци и MIMIC-III Waveform Matched Subset база на податоци.

Базата на податоци MIMIC-III Waveform содржи илјадници записи на повеќе физиолошки сигнали („бранови форми“) и временски низи на витални знаци („нумерички“) собрани со мониторинг на пациенти (возрасни и новороденчиња во кревет) сместени на одделот за интензивна нега. Всушност, оваа база содржи 67830 сетови на записи за околу 30000 пациенти. Скоро сите сетови на записи вклучуваат записи за бранови кои содржат дигитализирани сигнали (обично ЕКГ, АВР, фреквенција на дишење и SpO₂, но и други сигнали) и „нумерички“ записи кои содржат временска серија на периодични мерења, при што секое претставува квазиконтинуиран запис на виталните знаци на поединечен пациент во текот на престојот на интензивна нега (неколку дена или неколку недели). Секој сет од записи е сместен во соодветен директориум со името на записот. Базата содржи десет директориуми од 30 до 39 и тие се модифицирани последен пат во 2017 година.

Придружник на оваа база е клиничката база на податоци MIMIC-III, која содржи детални клинички информации за многу од пациентите претставени во Waveform базата на податоци. Податоците вклучуваат витални знаци, лекови, лабораториски мерења, набљудувања и белешки креирани од давателите на здравствената заштита, биланс на течности, кодови на процедури, дијагностички кодови, извештаи со слики, должина на престој во болница, податоци за преживување и друго. Содржи клинички записи за над на 40000 субјекти. MIMIC-III Waveform Matched Subset базата на податоци содржи 22317 записи за брановите и 22247 нумерички записи, кои се исти и временски усогласени со 10282 записи од MIMIC-III клиничката база на податоци.

Отворените податоци од клиничките испитувања кои се наоѓаат во базата PPG-BP, даваат можност на истражувачите ширум светот да постават нови хипотези, да ги потврдат објавените резултати и да постигнат нови научни достигнувања во медицинските истражувања. Оваа здравствена база на податоци која располага со податоци за неинвазивно откривање на кардиоваскуларни

заболувања (CVD), содржи 657 записи со податоци од вкупно 219 испитаници. Базата на податоци содржи податоци за пациенти од 20 до 89-годишна возраст, како и евиденција на нивните заболувања, вклучувајќи хипертензија и дијабетес. Собирањето на податоците било извршено под контрола на стандардни експериментални услови и спецификации. Оваа база на податоци може да се користи за да се спроведе студија за евалуација на квалитетот на сигналот добиен од фотоплетизмограф (PPG) и да се истражи вродената врска помеѓу PPG брановата форма и кардиоваскуларните заболувања за да се откријат и проценат латентните карактеристични информации содржани во PPG сигналите. Овие податоци, исто така, можат да се користат за проучување на рано и неинвазивно контролно тестирање на вообичаени кардиоваскуларни заболувања како што се хипертензија и други сродни заболувања на кардиоваскуларниот систем, како што е дијабетесот. Податоците во оваа база датираат од 2018 година.

Од оваа база искористивме околу 1500 податоци кои содржат информации за PPG сигнали, а истовремено користевме и податоци за крвниот притисок. Овие податоци ги користиме како тренинг податоци во процесот на машинско учење.

5.3 Класификација со помош на околината WEKA

Како алатка за испробување на методите за машинско учење, кои ги опишавме во претходното поглавје ја користиме околината WEKA.

Кратенката WEKA е добиена од сложенката Waikato Enviroment for Knowledge Analysis, што во превод значи Waikato околина за анализа на знаење. Системот WEKA е напишан во програмскиот јазик JAVA. WEKA е софтвер за машинско учење/софтвер за податочно рударење кој се користи за, образование и апликации.

Постојат неколку нивоа на користење на WEKA системот. Прво, направена е имплементација на алгоритми за учење што можат да се применуваат врз податоците за тренинг директно од командна линија. Исто така, вклучено е множество од алатки за трансформација на множества од податоци, како што се алгоритми за дискретизација и различни типови на филтри за податоците. Еден начин на користење на WEKA е да се примени методот за учење врз множеството

податоци и да се анализира неговиот излез и да се екстрахираат информациите за податоците. Друг начин е да се применат неколку алгоритми за учење и да се споредат нивните перформанси за да се избере најдобриот за предвидување. Методите за учење се наречени класификатори. Сите тие алгоритми имаат ист интерфејс од командна линија и постои множество од генерички опции од командна линија и опции кои се карактеристични за секој алгоритам посебно.

Перформансите на сите класификатори се мерат со ист модул за евалуација. Освен алгоритми за учење, WEKA обезбедува и филтри за пред- и постпроцесирање на податоците. Како и кај класификаторите и филтрите имаат стандардизиран интерфејс од командна линија. Основната цел на WEKA е развивање на алгоритми за класификација и филтрирање на податоци.

Во WEKA, освен овие алгоритми се вклучени и алгоритми за учење на асоцијациски правила и кластерирање на податоци.

Главни карактеристики на WEKA се:

- Големо множество на податоци со алатки за обработка, алгоритми за учење и методи на евалуација.
- Графички кориснички интерфејси (вклучувајќи и визуелизација на податоци).
- Околина за споредување на алгоритми за учење.

Во WEKA се имплементирани следните методи: Bayes, Decision trees and rules, Neural networks, Functions, Meta, lazy classifiers итн. со следните опции за испитување: користење множество за обука (тренинг) (Use training set), обезбедено тест-множество (Supplied test set), крос-валидација (Cross-validation), процентуална поделба (Percentage split).

- Use training set класификаторот се евалуира според тоа колку добро врши предвидување на класите примери на кои се тренира. Бројот на погрешни класификации на податоци за тренирање не е добар показател на перформансите на предикативниот модел.
- Supplied test set класификаторот се евалуира според тоа колку добро го врши предвидувањето на класата според примерите кои се вчитуваат од датотеката.

- Cross-validation класификаторот се евалуира со помош на вкрстено потврдување користејќи различен број на групи. Прво, сите податоци се поделуваат на k подгрупи податоци со еднаква далечина. Сите тие групи служат како податоци за тренирање на моделот. Тоа се вика k -foldcross-validation. Поделбата на 10 делови се покажува како најдобра во пракса.
- Percentage split, класификаторот се евалуира според тоа колку добро врши предвидување на податоците кои се користат, односно се задржуваат за тестирање.

Главната разлика помеѓу сетот за обука и cross validation (крос-валидацијата) е во следново: кога користиме сет за обука, тогаш ако имаме 100 примероци Weka користи класификациски алгоритам кој сме го избрале за да изгради модел од сите 100 примероци. Потоа за да го изврши тестирањето ги користи сите 100 примероци. Но, кога користиме, на пример, крос-валидација 10, тогаш од 100 примероци се прават 10 различни модели со по 10 примероци и добиваме некое средно време на извршување. Понекогаш е потребно да се користи сет за обука, но најчесто најдобри и највалидни резултати се добиваат со користење на крос-валидација 10.

Учењето на дрва за одлучување во системот WEKA е имплементирано преку множество од класи.

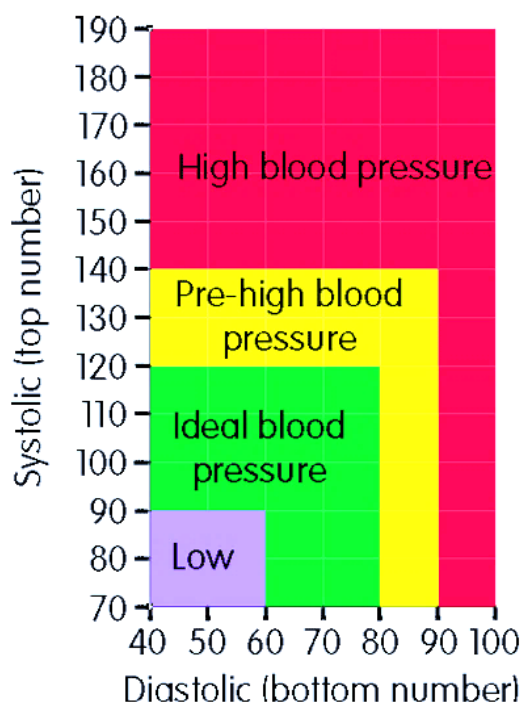
Секоја класа посебно има определена улога во алгоритамот. Концептуално гледано, класите можат да се поделат во неколку групи според функцијата што ја имаат во алгоритамот и тоа: класи за манипулација со дрва за одлучување, класи за селекција на модел на дрво, класи за поделба на ниво на јазол, класи за критериуми за поделба, класа за пресметка на распределбата на веројатност на ниво на јазол и друго.

WEKA работи со фајлови во ARFF формат (атрибут — врска формат на датотеката или Attribute-Relation File Format) кој е ASCII текстуална датотека која опишува листа на примероци кои делат заедничко множество на атрибути.

Ние ја користиме оваа алатка поради потребите од користење на машинското учење врз нашите податоци.

5.4 Резултати од класификација на податоците

Бидејќи целта е да го предвидиме крвниот притисок, како излезе ги користиме податоците за систолниот и дијастолниот дел од крвниот притисок, а како влезови сите морфолошки карактеристики на PPG-сигналот, претходно опишани (во табела 5. 2). Вредностите на систолниот и дијастолниот дел можат многу да варираат па за тоа ги делиме во групи, слика5.8.



Слика 5.8: Вредности за крвен притисок кај возрасни

Figure 5.8 Blood pressure values for adults

За дијастолниот дел од крвниот притисок, групите ги означуваме со вредности од 0-3

- 0 - дијастолен крвен притисок со вредност помала (или до) 60 (ниска вредност)
- 1 - дијастолен крвен притисок во интервал од 61 до 80 (нормална вредност)
- 2 - дијастолен крвен притисок во интервал од 81 до 90 (предвисока вредност)

- 3 - дијастолен крвен притисок со вредност поголема од 90 (висока вредност).

За систолниот дел од крвниот притисок поделбата по групи е следна:

- 0 - систолен крвен притисок со вредност помала (или до) 90 (ниска вредност)
- 1 - систолен крвен притисок во интервал од 91 до 120 (нормална вредност)
- 2 - систолен крвен притисок во интервал од 121 до 140 (предвисока вредност)
- 3 - систолен крвен притисок со вредност поголема од 141 (висока вредност).

Од методите за машинско учење, ние со нашите податоци ги користиме алгоритмите за класификација, имплементирани во WEKA: J48, Random Tree, Random Forest, LMT, Bayes Net, SMO, Multi-Layer Perceptron, KStar, AdaBoostM1 и JRip.

Табела 5.3: Резултати за процент на точно класифицирани податоци од алгоритмите за класификација во WEKA

Table 5.3 Results for percentage of accurately classified data from WEKA classification algorithms

	SBP	DBP
Алгоритам	<i>Процент на точно класифицирани податоци</i>	
J48	78.0714 %	81.5 %
RandomTree	94.2143 %	95 %
RandomForest	96.1429 %	95.6429 %
LMT	92.7857 %	88.8571 %
BayesNet	53.5714 %	62.5714 %
SMO	53.1429 %	58.5 %
Multi-LayerPerceptron	64.3571 %	72.7857 %
KStar	95.7143 %	97 %
AdaBoostM1	49.1429 %	56.2857 %
JRip	65.1429 %	73.4286 %

Од резултатите прикажани на табела 5.3, се гледа дека најдобри резултати се добиваат со користење на алгоритмите за класифицирање базирани на дрва, а посебно алгоритмот за Random Forest (96.1429 % за SBP и 95.6429 % за DBP). Исто така и алгоритмот KStar покажува голем процент на точно класифицирани податоци и кај систолниот и кај дијастолниот крвен притисок (95.7143 % и 97 % соодветно).

Ова е случај кога како влезови се користат сите карактеристики прикажани на табела 5. 2. Но, како што споменавме и претходно, во многу случаи тешко е определувањето на секундарниот врв, односно определувањето на точките или индексите на *dicrotic notch* и *dicrotic peak*. Тоа е многу честа појава, бидејќи PPG-сигналот е многу подложен на шумови, па во тој случај, многу тешко е определувањето и на главните минимуми и максимуми а речиси невозможно е определувањето на *dicrotic notch* и *dicrotic* врвот. Во таков случај, пожелно е да се користат само индексите на минимумите и максимумите во PPG-сигналот, кои како што и претходно објаснивме, ги добиваме со помош на AMPD-алгоритмот. На тој начин се намалува бројот на влезни параметри, односно бројот на морфолошките карактеристики на бранот кои ги користиме како влезови во алгоритмите за машинско учење. Карактеристиките кои во таков случај ги користиме се: T , T_s , T_d , AUC_{sys} , AUC_{dia} и H . Кога со помалку влезни параметри се извршуваат алгоритмите за класификација, времето за нивна обработка се намалува, но очекувано, се намалува и процентот на точно класифицираните податоци. Резултатите кои ги добиваме со намалениот број на влезни параметри се дадени во табела 5.4.

Табела 5.4: Резултати за процент на точно класифицирани податоци од алгоритмите за класификација во WEKA, со помал број влезови

Table 5.4 Results for percentage of accurately classified data from WEKA classification algorithms, with a smaller number of inputs

	SBP	DBP
Алгоритам	Процент на точно класифицирани податоци	
J48	76.5 %	78 %

RandomTree	93.3571 %	95.1429 %
RandomForest	96.0714 %	96.2857 %
LMT	89 %	84 %
BayesNet	52.2143 %	60.7857 %
SMO	53.0714 %	58.3571 %
Multi-LayerPerceptron	58.5714 %	68.4286 %
KStar	95.1429 %	96 %
AdaBoostM1	49.1429 %	56.2857 %
JRip	63.6429 %	70.0714 %

Процентот на точно класифицирани инстанци, во нашиот случај со намален број на влезните параметри, сепак останува висок кај Random Forest (вредности: 96.0714 %за SBP и 96.2857 % за DBP) и KStar (вредности: 95.1429 %за SBP и 96 %за DBP).

6. Заклучок

Со напредокот во обработката на дигиталните сигнали и зголемената компјутерска моќ, а исто така и со зголемената потреба за производи за следење на здравјето, значително внимание добиваат компјутерската автоматска естимација на биофизиолошките параметри, особено со постоењето на нови сензори што лесно се носат и можат да се искористат за континуирани неинвазивни решенија за мерење на ваквите параметри. Овие трендови исто така поттикнуваат развој на нови методи и системи за добивање на што попрецизни резултати.

Ние во овој труд се задржавме на крвниот притисок, како биофизиолошки параметар кој има клучно значење за здравјето на човекот, а со негово континуирано следење може да се превенира и да се алармира за многу здравствени проблеми. Но истражувањето би можело да се прошири и на другите биофизиолошки параметри.

Според разгледаните методологии, сигналите ЕКГ и PPG заедно се најкористениот начин за проценка на крвниот притисок користејќи методи на регресија или во последните студии, вклучувајќи невронски мрежи. Но, користењето на ЕКГ и PPG сигналите заедно не е многу погодно затоа што се потребни два различни бранови и со тоа повеќе сензори. Со напредокот на системите интегрирани во чип, можат да се интегрираат дури и моќни единици за обработка на графика, а со тоа се нудат можности за обработка на сигнали во реално време и извршување на сложени алгоритми. Затоа се засилуваат истражувањата за континуирана и точна естимација на крвен притисок директно и само од сигналите на ЕКГ или само од PPG сигналите. Примената на методите за машинско учење во овој контекст се смета за ветувачки пристап.

Развиен е прототип на систем за континуирано следење на биофизиолошкиот параметар за крвен притисок. Овој систем користи иновативна технологија која работи на принципот на пресметка во магла. Целта на ваквиот систем е да овозможи навремено алармирање кога ќе настанат некои брзи промени или флукутации во биофизиолошкиот параметар, а со тоа да се обезбеди

соодветна навремена реакција. На тој начин би можело да се спречат многу сериозни здравствени ситуации кои можат да бидат живото загрижувачки или со тешки последици.

Бидејќи за ваквиот систем, потребно навремено алармирање и работа во реално време, ние предложивме нова архитектура. Во оваа архитектура која се базира на пресметка во магла, податоците се претпроцесираат во мрежните уреди поставени на рабовите на мрежата. На таков начин се подобрува ефикасноста и се зголемуваат перформансите на самиот систем, а од друга страна истовремено се намалува мрежниот сообраќај, а со тоа и латентноста.

Системот е составен од сензори кои многу лесно можат да се носат во текот на целиот ден и не претставуваат никаква пречка во извршувањето на секојдневните дејствија.

Информациите собрани од сензорите понатаму се обработуваат за да се извлечат потребните биофизиолошки параметри. За таа цел се имплементирани алгоритми кои се засноваат на методите на машинско учење, што резултира во соодветен одговор во реално време.

Ние во овој труд се задржавме на естимација на крвен притисок со помош и единствено со PPG сигналите, бидејќи овој пристап е најмалку истражуван, а од друга страна PPG сензорите се лесни за носење и не претставуваат пречка во извршувањето на секојдневните активности, а истовремено се и евтини и лесно достапни. Ваквата техника, за пресметка на крвниот притисок со користење само на PPG-сигналот, ние ја искористивме за развој на преносен ненаметлив систем и систем во реално време, способен за постојано следење на пациентите за долг временски период и во динамични услови.

За потребите на системот развивме и алгоритам за естимација на PPG-сигналот. Но, овој сигнал е многу осетлив на шум од надворешната средина и посебно шум добиен при движењето и извршувањето на секојдневните дејствија, прво PPG-сигналот го прочистуваме со предложен алгоритам, а потоа врз основа на чистиот сигнал правиме естимација на крвниот притисок.

Користиме и методи на машинско учење за регресија и класификација за да добиеме предвидување за крвниот притисок врз база на извлечените

карактеристики од PPG-сигналот. Според резултатите кои ги добивме, најпрецизни предвидувања се добиваат со методот на случајни шуми (Random Forest).

Бидејќи во овој труд се задржуваме само на PPG-сигналот и естимација за крвниот притисок со него, нашата цел за идни истражувања ќе бидат во насока на проширување на повеќе биофизиолошките параметри со што би се добила уште попрецизна слика за здравствената состојба на човекот.

Ваквиот систем за континуирано следење на биофизиолошките параметри е особено применлив во системите за амбиентално помогнато живеење како што се домовите за нега на стари лица, но и во други медицински установи, во кои со користење на ваквите технологии може да се помогне и да се олесни процесот на следење на здравствената состојба на пациентите, а од друга страна им се нуди зголемена удобност и безбедност на пациентите и точна и прецизна информираност.

Затоа, дел од нашите понатамошни истражувања треба да е насочен кон доразвивање и усовршување на ваков систем и негова апликација во установи кај кои постојаното следење на вакви биофизиолошки параметри е од витално значење за зачувување на здравствената состојба и продолжување на квалитетен животот.

Користена литература

- Aarts, V., Dellimore, K. H., Wijshoff, R., Derkx, R., Van De Laar, J., & Muehlsteff, J. (2017). Performance of an accelerometer-based pulse presence detection approach compared to a reference sensor. *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2017*. <https://doi.org/10.1109/BSN.2017.7936033>
- Aazam, M., & Huh, E. N. (2015). Fog computing micro datacenter based dynamic resource estimation and pricing model for IoT. *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*. <https://doi.org/10.1109/AINA.2015.254>
- Achkoski, J., Koceski, S., Bogatinov, D., Temelkovski, B., Stevanovski, G., & Kocev, I. (2017). Remote triage support algorithm based on fuzzy logic. *Journal of the Royal Army Medical Corps*. <https://doi.org/10.1136/jramc-2015-000616>
- Addison, P. S., Watson, J. N., Mestek, M. L., Ochs, J. P., Uribe, A. A., & Bergese, S. D. (2014). Pulse oximetry-derived respiratory rate in general care floor patients. *Journal of Clinical Monitoring and Computing*. <https://doi.org/10.1007/s10877-014-9575-5>
- Ahmad, S., Chen, S., Soueidan, K., Batkin, I., Bolic, M., Dajani, H., & Groza, V. (2012). Electrocardiogram-assisted blood pressure estimation. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2011.2180019>
- Ahrens, T. (2008). The most important vital signs are not being measured. In *Australian Critical Care*. <https://doi.org/10.1016/j.aucc.2007.12.061>
- Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. In *Physiological Measurement*. <https://doi.org/10.1088/0967-3334/28/3/R01>
- AlMahamdy, M., & Riley, H. B. (2014). Performance study of different denoising methods for ECG signals. *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2014.08.048>
- Appelboom, G., Camacho, E., Abraham, M. E., Bruce, S. S., Dumont, E. L. P., Zacharia, B. E., D'Amico, R., Slomian, J., Reginster, J. Y., Bruyère, O., & Connolly, E. S. (2014). Smart wearable body sensors for patient self-assessment and monitoring. In *Archives of Public Health*. <https://doi.org/10.1186/2049-3258-72-28>

- Axak, N., Rosinskiy, D., Barkovska, O., & Novoseltsev, I. (2018). Cloud-fog-dew architecture for personalized service-oriented systems. *Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT 2018*. <https://doi.org/10.1109/DESSERT.2018.8409103>
- Baca, A., Biagetti, G., Camilletti, M., Crippa, P., Falaschetti, L., Orcioni, S., Rossini, L., Tonelli, D., & Turchetti, C. (2015). CARMA: A robust motion artifact reduction algorithm for heart rate monitoring from PPG signals. *2015 23rd European Signal Processing Conference, EUSIPCO 2015*. <https://doi.org/10.1109/EUSIPCO.2015.7362864>
- Baek, H. J., Lee, H. B., Kim, J. S., Choi, J. M., Kim, K. K., & Park, K. S. (2009). Nonintrusive biological signal monitoring in a car to evaluate a driver's stress and health state. *Telemedicine and E-Health*. <https://doi.org/10.1089/tmj.2008.0090>
- Bahl, P., Han, R. Y., Li, L. E., & Satyanarayanan, M. (2012). Advancing the state of mobile cloud computing. *MCS'12 - Proceedings of the 3rd ACM Workshop on Mobile Cloud Computing and Services*. <https://doi.org/10.1145/2307849.2307856>
- Banaee, H., Ahmed, M. U., & Loutfi, A. (2013). Data mining for wearable sensors in health monitoring systems: A review of recent trends and challenges. In *Sensors (Switzerland)*. <https://doi.org/10.3390/s131217472>
- Bandodkar, A. J., & Wang, J. (2014). Non-invasive wearable electrochemical sensors: A review. In *Trends in Biotechnology*. <https://doi.org/10.1016/j.tibtech.2014.04.005>
- Beck, M. T., Werner, M., Feld, S., & Schimper, T. (2014). Mobile Edge Computing : A Taxonomy. *Proc. of the Sixth International Conference on Advances in Future Internet*.
- Billard, L., & Diday, E. (2002). *Symbolic Regression Analysis*. https://doi.org/10.1007/978-3-642-56181-8_31
- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. In *Information Science and Statistics*.
- Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012a). Fog computing and its role in the internet of things. *MCC'12 - Proceedings of the 1st ACM Mobile Cloud Computing Workshop*. <https://doi.org/10.1145/2342509.2342513>
- Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012b). Fog Computing and Its Role in

- the Internet of Things Characterization of Fog Computing. *MCC'12 Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*.
- Breiman, L. (2001). ST4_Method_Random_Forest. *Machine Learning*.
<https://doi.org/10.1017/CBO9781107415324.004>
- Breiman, L. (1996). Bagging predictors - Springer. *Machine Learning*.
<https://doi.org/10.1007/BF00058655>
- Breiman, Leo, Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. In *Classification and Regression Trees*.
<https://doi.org/10.1201/9781315139470>
- Buller, M. J., Tharion, W. J., Hoyt, R. W., & Chadwicke Jenkins, O. (2010). Estimation of human internal temperature from wearable physiological sensors. *Proceedings of the National Conference on Artificial Intelligence*.
- Buxi, D., Redoute, J. M., & Yuce, M. R. (2015). Cuffless blood pressure estimation from the carotid pulse arrival time using continuous wave radar. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. <https://doi.org/10.1109/EMBC.2015.7319687>
- Casson, A. J., Vazquez Galvez, A., & Jarchi, D. (2016). Gyroscope vs. accelerometer measurements of motion from wrist PPG during physical exercise. *ICT Express*.
<https://doi.org/10.1016/j.icte.2016.11.003>
- Cearley, D. W. (2010). Cloud Computing: Key Initiative Overview. In *Analysis*.
- Center for Disease Control and Prevention. (2011). Prescription painkiller overdoses at epidemic levels. *U.S. Department of Health and Human Resources*.
- Chakravorty, R. (2006). A programmable service architecture for mobile medical care. *Proceedings - Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2006*.
<https://doi.org/10.1109/PERCOMW.2006.11>
- Chan, M., Estève, D., Fourniols, J. Y., Escriba, C., & Campo, E. (2012). Smart wearable systems: Current status and future challenges. *Artificial Intelligence in Medicine*.
<https://doi.org/10.1016/j.artmed.2012.09.003>
- Chandrasekaran, V., Dantu, R., Jonnada, S., Thiyagaraja, S., & Subbu, K. P. (2013). Cuffless differential blood pressure estimation using smart phones. *IEEE Transactions*

- on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2012.2211078>
- Chen, S., Ji, Z., Wu, H., & Xu, Y. (2019). A non-invasive continuous blood pressure estimation approach based on machine learning. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19112585>
- Chen, Y., Wen, C., Tao, G., & Bi, M. (2012). Continuous and noninvasive measurement of systolic and diastolic blood pressure by one mathematical model with the same model parameters and two separate pulse wave velocities. *Annals of Biomedical Engineering*. <https://doi.org/10.1007/s10439-011-0467-2>
- Chiu, Y. Y., Lin, W. Y., Wang, H. Y., Huang, S. Bin, & Wu, M. H. (2013). Development of a piezoelectric polyvinylidene fluoride polymer-based sensor patch for simultaneous heartbeat and respiration monitoring. *8th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems, IEEE NEMS 2013*. <https://doi.org/10.1109/NEMS.2013.6559679>
- Chong, J. W., Dao, D. K., Salehizadeh, S. M. A., McManus, D. D., Darling, C. E., Chon, K. H., & Mendelson, Y. (2014). Photoplethysmograph Signal Reconstruction Based on a Novel Hybrid Motion Artifact Detection–Reduction Approach. Part I: Motion and Noise Artifact Detection. *Annals of Biomedical Engineering*. <https://doi.org/10.1007/s10439-014-1080-y>
- Chowdhury, S. S., Hyder, R., Hafiz, M. S. Bin, & Haque, M. A. (2018). Real-Time Robust Heart Rate Estimation from Wrist-Type PPG Signals Using Multiple Reference Adaptive Noise Cancellation. *IEEE Journal of Biomedical and Health Informatics*. <https://doi.org/10.1109/JBHI.2016.2632201>
- Chung, E., Chen, G., Alexander, B., & Cannesson, M. (2013). Non-invasive continuous blood pressure monitoring: A review of current applications. In *Frontiers of Medicine in China*. <https://doi.org/10.1007/s11684-013-0239-5>
- Cisco. (2015). Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things Connect. *White Paper*.
- Cohen, W. W. (1995). Fast Effective Rule Induction. In *Machine Learning Proceedings 1995*. <https://doi.org/10.1016/b978-1-55860-377-6.50023-2>
- Dastjerdi, A. V., Gupta, H., Calheiros, R. N., Ghosh, S. K., & Buyya, R. (2016). Fog Computing: Principles, architectures, and applications. In *Internet of Things: Principles*

- and Paradigms*. <https://doi.org/10.1016/B978-0-12-805395-9.00004-6>
- Dias, D., & Cunha, J. P. S. (2018). Wearable health devices—vital sign monitoring, systems and technologies. In *Sensors (Switzerland)*. <https://doi.org/10.3390/s18082414>
- Dias, D., Ferreira, N., & Cunha, J. P. S. (2017). VitalLogger: An adaptable wearable physiology and body-Area ambiance data logger for mobile applications. *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2017*. <https://doi.org/10.1109/BSN.2017.7936010>
- Dohr, A., Modre-Osprian, R., Drobits, M., Hayn, D., & Schreier, G. (2010). The internet of things for ambient assisted living. *ITNG2010 - 7th International Conference on Information Technology: New Generations*. <https://doi.org/10.1109/ITNG.2010.104>
- Durrant-Whyte, H. F. (1988). Sensor Models and Multisensor Integration. *The International Journal of Robotics Research*. <https://doi.org/10.1177/027836498800700608>
- Elgendi, M., Norton, I., Brearley, M., Abbott, D., & Schuurmans, D. (2014). Detection of a and b waves in the acceleration photoplethysmogram. *BioMedical Engineering Online*. <https://doi.org/10.1186/1475-925X-13-139>
- Elliott, M., & Coventry, A. (2012). Critical care: The eight vital signs of patient monitoring. *British Journal of Nursing*. <https://doi.org/10.12968/bjon.2012.21.10.621>
- Evans. (2011). How the Next Evolution of the Internet Is Changing Everything, White paper, 2011. In *CISCO white paper*.
- Fallet, S., & Vesin, J. M. (2017). Robust heart rate estimation using wrist-type photoplethysmographic signals during physical exercise: An approach based on adaptive filtering. *Physiological Measurement*. <https://doi.org/10.1088/1361-6579/aa506e>
- Farahani, B., Barzegari, M., Shams Aliee, F., & Shaik, K. A. (2020). Towards collaborative intelligent IoT eHealth: From device to fog, and cloud. *Microprocessors and Microsystems*. <https://doi.org/10.1016/j.micpro.2019.102938>
- Farahani, B., Firouzi, F., Chang, V., Badaroglu, M., Constant, N., & Mankodiya, K. (2018). Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare. *Future Generation Computer Systems*.

<https://doi.org/10.1016/j.future.2017.04.036>

- Fujita, Y., Hiromoto, M., & Sato, T. (2018). PARHELIA: Particle filter-based heart rate estimation from photoplethysmographic signals during physical exercise. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2017.2697911>
- Fürnkranz, J., & Widmer, G. (1994). Incremental Reduced Error Pruning. In *Machine Learning Proceedings 1994*. <https://doi.org/10.1016/b978-1-55860-335-6.50017-9>
- Gao, M., Olivier, N. B., & Mukkamala, R. (2016). Comparison of noninvasive pulse transit time estimates as markers of blood pressure using invasive pulse transit time measurements as a reference. *Physiological Reports*. <https://doi.org/10.14814/phy2.12768>
- Garcia, A., Rack-Gomer, A. L., Bhavaraju, N. C., Hampapuram, H., Kamath, A., Peyser, T., Facchinetti, A., Zecchin, C., Sparacino, G., & Cobelli, C. (2013). Dexcom G4AP: An advanced continuous glucose monitor for the artificial pancreas. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/193229681300700604>
- Gaura, E., Kemp, J., & Brusey, J. (2013). Leveraging knowledge from physiological data: On-body heat stress risk prediction with sensor networks. *IEEE Transactions on Biomedical Circuits and Systems*. <https://doi.org/10.1109/TBCAS.2013.2254485>
- Gesche, H., Grosskurth, D., Kuchler, G., & Patzak, A. (2012). Continuous blood pressure measurement by using the pulse transit time: Comparison to a cuff-based method. *European Journal of Applied Physiology*. <https://doi.org/10.1007/s00421-011-1983-3>
- Gia, T. N., Jiang, M., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2015). Fog computing in healthcare Internet of Things: A case study on EKG feature extraction. *Proceedings - 15th IEEE International Conference on Computer and Information Technology, CIT 2015, 14th IEEE International Conference on Ubiquitous Computing and Communications, IUCC 2015, 13th IEEE International Conference on Dependable, Autonomic and Se.* <https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.51>
- Gia, T. N., Thanigaivelan, N. K., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2015). Customizing 6LoWPAN networks towards Internet-of-Things based ubiquitous healthcare systems. *NORCHIP 2014 - 32nd NORCHIP Conference:*

- The Nordic Microelectronics Event*. <https://doi.org/10.1109/NORCHIP.2014.7004716>
- Giang, N. K., Blackstock, M., Lea, R., & Leung, V. C. M. (2015). Developing IoT applications in the Fog: A Distributed Dataflow approach. *Proceedings - 2015 5th International Conference on the Internet of Things, IoT 2015*. <https://doi.org/10.1109/IOT.2015.7356560>
- Giovangrandi, L., Inan, O. T., Banerjee, D., & Kovacs, G. T. A. (2012). Preliminary results from BCG and EKG measurements in the heart failure clinic. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. <https://doi.org/10.1109/EMBC.2012.6346790>
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C., & Stanley, H. E. (2000). Components of a New Research Resource for Complex Physiologic Signals. *Circulation*.
- Goli, S., & Jayanthi, T. (2014). Cuff less Continuous Non-Invasive Blood Pressure Measurement Using Pulse Transit Time Measurement. *International Journal of Recent Development in Engineering and Technology*.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning - An MIT Press book*. MIT Press.
- Greenfield, D. (2016). *Fog Computing vs. Edge Computing: What's the Difference?* | *Automation World*. Automation World.
- Gu, W. B., Poon, C. C. Y., Leung, H. K., Sy, M. Y., Wong, M. Y. M., & Zhang, Y. T. (2009). A novel method for the contactless and continuous measurement of arterial blood pressure on a sleeping bed. *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*. <https://doi.org/10.1109/IEMBS.2009.5335393>
- Guo, li, Berglin, L., Wiklund, U., & Mattila, H. (2013). Design of a garment-based sensing system for breathing monitoring. *Textile Research Journal*. <https://doi.org/10.1177/0040517512444336>
- Haldar, R. N. (2013). Global Brief on Hypertension: Silent Killer, Global Public Health Crisis. *Indian Journal of Physical Medicine and Rehabilitation*. <https://doi.org/10.5005/ijopmr-24-1-2>
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. In *Data*

- Mining: Concepts and Techniques*. <https://doi.org/10.1016/C2009-0-61819-5>
- Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*. [https://doi.org/10.6339/JDS.2007.05\(2\).396](https://doi.org/10.6339/JDS.2007.05(2).396)
- Hertzman, A. B. (1938). THE BLOOD SUPPLY OF VARIOUS SKIN AREAS AS ESTIMATED BY THE PHOTOELECTRIC PLETHYSMOGRAPH. *American Journal of Physiology-Legacy Content*. <https://doi.org/10.1152/ajplegacy.1938.124.2.328>
- Iorga, M., Martin, M. J., & Feldman, L. (2018). Fog Computing Conceptual Model NIST Special Publication 500-325. *National Institute of Standards and Technology*.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning, Springer Texts. In *Springer Texts*. <https://doi.org/10.1016/j.peva.2007.06.006>
- Jarchi, D., & Casson, A. J. (2017). Towards Photoplethysmography-Based Estimation of Instantaneous Heart Rate during Physical Activity. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2017.2668763>
- Jin, A., Yin, B., Morren, G., Duric, H., & Aarts, R. M. (2009). Performance evaluation of a tri-axial accelerometry-based respiration monitoring for ambient assisted living. *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*. <https://doi.org/10.1109/IEMBS.2009.5333116>
- Johnston, W. S., Branche, P. C., Pujary, C. J., & Mendelson, Y. (2004). Effects of motion artifacts on helmet-mounted pulse oximeter sensors. *Bioengineering, Proceedings of the Northeast Conference*. <https://doi.org/10.1109/nebc.2004.1300071>
- Joseph, G., Joseph, A., Titus, G., Thomas, R. M., & Jose, D. (2014). Photoplethysmogram (PPG) signal analysis and wavelet de-noising. *2014 Annual International Conference on Emerging Research Areas: Magnetism, Machines and Drives, AICERA/ICMMD 2014 - Proceedings*. <https://doi.org/10.1109/AICERA.2014.6908199>
- Kachuee, Mohamad, Kiani, M. M., Mohammadzade, H., & Shabany, M. (2015). Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. *Proceedings - IEEE International Symposium on Circuits and Systems*. <https://doi.org/10.1109/ISCAS.2015.7168806>

- Kachuee, Mohammad, Kiani, M. M., Mohammadzade, H., & Shabany, M. (2017). Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2016.2580904>
- Kambourakis, G., Klaoudatou, E., & Gritzalis, S. (2007). Securing medical sensor environments: The CodeBlue framework case. *Proceedings - Second International Conference on Availability, Reliability and Security, ARES 2007*. <https://doi.org/10.1109/ARES.2007.135>
- Kamshilin, A. A., & Margaryants, N. B. (2017). Origin of Photoplethysmographic Waveform at Green Light. *Physics Procedia*. <https://doi.org/10.1016/j.phpro.2017.01.024>
- Khalid, S. G., Zhang, J., Chen, F., & Zheng, D. (2018). Blood Pressure Estimation Using Photoplethysmography Only: Comparison between Different Machine Learning Approaches. *Journal of Healthcare Engineering*. <https://doi.org/10.1155/2018/1548647>
- Khan, E., Al Hossain, F., Uddin, S. Z., Alam, S. K., & Hasan, M. K. (2016). A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2015.2466075>
- Kim, B. S., & Yoo, S. K. (2006). Motion artifact reduction in photoplethysmography using independent component analysis. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2005.869784>
- Kim, Jeehoon, Kwon, S., Seo, S., & Park, K. (2014). Highly wearable galvanic skin response sensor using flexible and conductive polymer foam. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. <https://doi.org/10.1109/EMBC.2014.6945148>
- Kim, Jungsoo, Park, J., Kim, K., Chee, Y., Lim, Y., & Park, K. (2007). Development of a nonintrusive blood pressure estimation system for computer users. *Telemedicine Journal and E-Health*. <https://doi.org/10.1089/tmj.2006.0034>
- Klabunde, R. E. (2012). Cardiovascular Physiology Concepts Second Edition. In *Lippincott Williams & Wilkins*.
- Klas, G. I. (2015). Fog Computing and Mobile Edge Cloud Gain Momentum Open Fog

- Consortium, ETSI MEC and Cloudlets. *White Paper*.
- Ko, J. G., Musăloiu-Elefteri, R., Lim, J. H., Chen, Y., Terzis, A., Gao, T., Destler, W., & Selavo, L. (2008). MEDiSN: Medical emergency detection in sensor networks. *SenSys'08 - Proceedings of the 6th ACM Conference on Embedded Networked Sensor Systems*. <https://doi.org/10.1145/1460412.1460452>
- Koceska, N., Koceski, S., Sazdovski, V., & Ciambone, D. (2017). Robotic assistant for elderly care – Development and evaluation. *International Journal of Automation Technology*. <https://doi.org/10.20965/ijat.2017.p0425>
- Koceska, N., Koceski, S., Zobel, P. B., Trajkovic, V., & Garcia, N. (2019). A telemedicine robot system for assisted and independent living. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19040834>
- Kocev, I., Achkoski, J., Bogatinov, D., Koceski, S., Trajkovic, V., Stevanoski, G., & Temelkovski, B. (2018). Novel approach for automating medical emergency protocol in military environment. *Technology and Health Care*. <https://doi.org/10.3233/THC-170852>
- Kodali, B. S. (2013). Capnography outside the operating rooms. In *Anesthesiology*. <https://doi.org/10.1097/ALN.0b013e318278c8b6>
- Koop, C., Mosher, R., Kun, L., Geiling, J., Grigg, E., Long, S., Macedonia, C., Merrell, R., Satava, R., & Rosen, J. (2008). Future delivery of health care: Cybercare. In *IEEE Engineering in Medicine and Biology Magazine*. <https://doi.org/10.1109/MEMB.2008.929888>
- Krehel, M., Schmid, M., Rossi, R. M., Boesel, L. F., Bona, G. L., & Scherer, L. J. (2014). An optical fibre-based sensor for respiratory monitoring. *Sensors (Switzerland)*. <https://doi.org/10.3390/s140713088>
- Krishnan, R., Natarajan, B., & Warren, S. (2010). Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2009.2039568>
- Kulev, I., Vlahu-Gjorgievska, E., Koceski, S., & Trajkovic, V. (2015). Evaluating an Ordered List of Recommended Physical Activities within Health Care System. *Advances in Intelligent Systems and Computing*. https://doi.org/10.1007/978-3-319-09879-1_24

- Kumar, N., Agrawal, A., & Deb, S. (2014). Cuffless BP measurement using a correlation study of pulse transient time and heart rate. *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*. <https://doi.org/10.1109/ICACCI.2014.6968642>
- Lameski, P., Zdravevski, E., Koceski, S., Kulakov, A., & Trajkovik, V. (2017). Suppression of Intensive Care Unit False Alarms Based on the Arterial Blood Pressure Signal. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2017.2690380>
- Laurin, A. & Joachim, J. Implementation of a feature detection algorithm for arterial blood pressure., 2017. URL <https://fr.mathworks.com/matlabcentral/fileexchange/60172-bp-annotate>.
- Le Gall, A., Laurin, A., Vallee, F., & Chemla, D. (2017). Comparison of systolic period duration using aortic flow or pressure based methods in anesthetized patients. *Computing in Cardiology*. <https://doi.org/10.22489/CinC.2017.324-370>
- Lee, C. M., & Zhang, Y. T. (2003). Reduction of motion artefacts from photoplethysmographic recordings using a wavelet denoising approach. *APBME 2003 - IEEE EMBS Asian-Pacific Conference on Biomedical Engineering 2003*. <https://doi.org/10.1109/APBME.2003.1302650>
- Lee, H., Chung, H., Ko, H., & Lee, J. (2018). Wearable Multichannel Photoplethysmography Framework for Heart Rate Monitoring during Intensive Exercise. *IEEE Sensors Journal*. <https://doi.org/10.1109/JSEN.2018.2801385>
- Lee, H., Chung, H., & Lee, J. (2019). Motion Artifact Cancellation in Wearable Photoplethysmography Using Gyroscope. *IEEE Sensors Journal*. <https://doi.org/10.1109/JSEN.2018.2879970>
- Li, P., Liu, M., Zhang, X., Hu, X., Pang, B., Yao, Z., & Chen, H. (2016). Novel wavelet neural network algorithm for continuous and noninvasive dynamic estimation of blood pressure from photoplethysmography. *Science China Information Sciences*. <https://doi.org/10.1007/s11432-015-5400-0>
- Liang, Y., Elgendi, M., Chen, Z., & Ward, R. (2018). Analysis: An optimal filter for short photoplethysmogram signals. *Scientific Data*. <https://doi.org/10.1038/sdata.2018.76>
- Lopez-Silva, S. M., Giannetti, R., Dotor, M. L., Silveira, J. P., Golmayo, D., Miguel-Tobal, F., Bilbao, A., Galindo, M., & Martin-Escudero, P. (2012). Heuristic algorithm

- for photoplethysmographic heart rate tracking during maximal exercise test. *Journal of Medical and Biological Engineering*. <https://doi.org/10.5405/jmbe.898>
- Luo, N., Ding, J., Zhao, N., Leung, B. H. K., & Poon, C. C. Y. (2014). Mobile health: Design of flexible and stretchable electrophysiological sensors for wearable healthcare systems. *Proceedings - 11th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2014*. <https://doi.org/10.1109/BSN.2014.25>
- Ma, H. T. (2014). A blood pressure monitoring method for stroke management. *BioMed Research International*. <https://doi.org/10.1155/2014/571623>
- Mahmood, Z. (2011). *Cloud Computing for Enterprise Architectures: Concepts, Principles and Approaches*. https://doi.org/10.1007/978-1-4471-2236-4_1
- Mahmood, Z., & Ramachandran, M. (2018). Fog computing: Concepts, principles and related paradigms. In *Fog Computing: Concepts, Frameworks and Technologies*. https://doi.org/10.1007/978-3-319-94890-4_1
- Mazeika, G. G., & Swanson, R. (2007). Respiratory Inductance Plethysmography: An Introduction. In *Pro-Tech Services*.
- McAdams, E., Krupaviciute, A., Gehin, C., Dittmar, A., Delhomme, G., Rubel, P., Fayn, J., & McLaughlin, J. (2011). Wearable electronic systems: Applications to medical diagnostics/monitoring. In *Wearable Monitoring Systems*. https://doi.org/10.1007/978-1-4419-7384-9_9
- McAdams, E. T., Gehin, C., Noury, N., Ramon, C., Nocua, R., Massot, B., Oliveira, A., Dittmar, A., Nugent, C. D., & McLaughlin, J. (2010). Biomedical sensors for ambient assisted living. *Lecture Notes in Electrical Engineering*. https://doi.org/10.1007/978-3-642-05167-8_14
- MedtronicDiabetes. (2015). Enlite Sensor. *Medtronic*.
- Mendelson, Y., Dao, D. K., & Chon, K. H. (2013). Multi-channel pulse oximetry for wearable physiological monitoring. *2013 IEEE International Conference on Body Sensor Networks, BSN 2013*. <https://doi.org/10.1109/BSN.2013.6575518>
- Mitchell, G. F. (2014). Arterial stiffness and hypertension. *Hypertension*. <https://doi.org/10.1161/HYPERTENSIONAHA.114.00921>
- Moosavi, S. R., Gia, T. N., Rahmani, A. M., Nigussie, E., Virtanen, S., Isoaho, J., & Tenhunen, H. (2015). SEA: A secure and efficient authentication and authorization

- architecture for IoT-based healthcare using smart gateways. *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2015.05.013>
- Mouradian, C., Naboulsi, D., Yangui, S., Glitho, R. H., Morrow, M. J., & Polakos, P. A. (2018). A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges. In *IEEE Communications Surveys and Tutorials*. <https://doi.org/10.1109/COMST.2017.2771153>
- Nagre, A., & Mendelson, Y. (2005). Effects of motion artifacts on pulse oximeter readings from different facial regions. *Bioengineering, Proceedings of the Northeast Conference*. <https://doi.org/10.1109/nebc.2005.1432000>
- Naha, R. K., Garg, S., Georgakopoulos, D., Jayaraman, P. P., Gao, L., Xiang, Y., & Ranjan, R. (2018). Fog computing: Survey of trends, architectures, requirements, and research directions. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2018.2866491>
- Nemati, E., Deen, M. J., & Mondal, T. (2012). A wireless wearable EKG sensor for long-term applications. *IEEE Communications Magazine*. <https://doi.org/10.1109/MCOM.2012.6122530>
- Nikolic-Popovic, J., & Goubran, R. (2011). Measuring heart rate, breathing rate and skin conductance during exercise. *MeMeA 2011 - 2011 IEEE International Symposium on Medical Measurements and Applications, Proceedings*. <https://doi.org/10.1109/MeMeA.2011.5966751>
- OpenFog Consortium Architecture Working Group. (2017). OpenFog Reference Architecture for Fog Computing. *OpenFog*. <https://doi.org/OPFRA001.020817>
- Otto, C., Milenković, A., Sanders, C., & Jovanov, E. (2006). System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of Mobile Multimedia*. <https://doi.org/10.1.1.77.2522>
- Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.1985.325532>
- Parák, J., & Havlík, J. (2011). EKG signal processing and heart rate frequency detection methods. In *Proceedings of Technical Computing Prague*.
- Pilt, K., Ferenets, R., Meigas, K., Lindberg, L. G., Temitski, K., & Viigimaa, M. (2013). New photoplethysmographic signal analysis algorithm for arterial stiffness estimation. *The Scientific World Journal*. <https://doi.org/10.1155/2013/169035>

- Quinlan, J. R. (1993). Combining Instance-Based and Model-Based Learning. In *Machine Learning Proceedings 1993*. <https://doi.org/10.1016/b978-1-55860-307-3.50037-x>
- R., M., & Prabhu, S. (2015). A Novel Algorithm to Obtain Respiratory Rate from the PPG Signal. *International Journal of Computer Applications*. <https://doi.org/10.5120/ijca2015906263>
- Rahmani, A. M., Gia, T. N., Negash, B., Anzanpour, A., Azimi, I., Jiang, M., & Liljeberg, P. (2018). Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. *Future Generation Computer Systems*. <https://doi.org/10.1016/j.future.2017.02.014>
- Ram, M. R., Madhav, K. V., Krishna, E. H., Komalla, N. R., & Reddy, K. A. (2012). A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Transactions on Instrumentation and Measurement*. <https://doi.org/10.1109/TIM.2011.2175832>
- Rastegar, S., Gholamhosseini, H., Lowe, A., Mehdipour, F., & Lindén, M. (2019). Estimating Systolic Blood Pressure Using Convolutional Neural Networks. *Studies in Health Technology and Informatics*.
- Reisner, A., Shaltis, P. A., McCombie, D., & Asada, H. H. (2008). Utility of the photoplethysmogram in circulatory monitoring. In *Anesthesiology*. <https://doi.org/10.1097/ALN.0b013e31816c89e1>
- Rosendorff, C., Lackland, D. T., Allison, M., Aronow, W. S., Black, H. R., Blumenthal, R. S., Cannon, C. P., De Lemos, J. A., Elliott, W. J., Findeiss, L., Gersh, B. J., Gore, J. M., Levy, D., Long, J. B., O'Connor, C. M., O'Gara, P. T., Ogedegbe, G., Oparil, S., & White, W. B. (2015). Treatment of Hypertension in Patients with Coronary Artery Disease: A Scientific Statement from the American Heart Association, American College of Cardiology, and American Society of Hypertension. *Hypertension*. <https://doi.org/10.1161/HYP.0000000000000018>
- Sanaei, Z., Abolfazli, S., Gani, A., & Buyya, R. (2014). Heterogeneity in mobile cloud computing: Taxonomy and open challenges. *IEEE Communications Surveys and Tutorials*. <https://doi.org/10.1109/SURV.2013.050113.00090>
- Sangeeta, B., & Laxmi, S. (2011). A Real Time Analysis of PPG Signal for

Measurement of SpO₂ and Pulse Rate. *International Journal of Computer Applications*.

Sannino, G., Melillo, P., Stranges, S., De Pietro, G., & Pecchia, L. (2015). Short term Heart Rate Variability to predict blood pressure drops due to standing: A pilot study. *BMC Medical Informatics and Decision Making*. <https://doi.org/10.1186/1472-6947-15-S3-S2>

Sannino, Giovanna, De Falco, I., & De Pietro, G. (2016). Genetic programming for a wearable approach to estimate blood pressure embedded in a mobile-based health system. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*. <https://doi.org/10.1109/ICTAI.2015.115>

Satyanarayanan, M., Bahl, P., Cáceres, R., & Davies, N. (2009). The case for VM-based cloudlets in mobile computing. *IEEE Pervasive Computing*. <https://doi.org/10.1109/MPRV.2009.82>

Satyanarayanan, M., Lewis, G., Morris, E., Simanta, S., Boleng, J., & Ha, K. (2013). The role of cloudlets in hostile environments. *IEEE Pervasive Computing*. <https://doi.org/10.1109/MPRV.2013.77>

Scholkmann, F., Boss, J., & Wolf, M. (2012). An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms*. <https://doi.org/10.3390/a5040588>

Seoane, F., Mohino-Herranz, I., Ferreira, J., Alvarez, L., Buendia, R., Ayllón, D., Llerena, C., & Gil-Pita, R. (2014). Wearable biomedical measurement systems for assessment of mental stress of combatants in real time. *Sensors (Switzerland)*. <https://doi.org/10.3390/s140407120>

Sharma, H., Sharma, K. K., & Bhagat, O. L. (2015). Respiratory rate extraction from single-lead EKG using homomorphic filtering. *Computers in Biology and Medicine*. <https://doi.org/10.1016/j.compbimed.2015.01.024>

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*. <https://doi.org/10.1109/JIOT.2016.2579198>

Silva, B. M. C., Rodrigues, J. J. P. C., Simoes, T. M. C., Sendra, S., & Lloret, J. (2014). An ambient assisted living framework for mobile environments. *2014 IEEE-EMBS*

- International Conference on Biomedical and Health Informatics, BHI 2014*.
<https://doi.org/10.1109/BHI.2014.6864399>
- Skala, K., Davidovic, D., Afgan, E., Sovic, I., & Sojat, Z. (2015). Scalable Distributed Computing Hierarchy: Cloud, Fog and Dew Computing. *Open Journal of Cloud Computing*.
- Slapničar, G., Mlakar, N., & Luštrek, M. (2019). Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19153420>
- Slapničar, G., Luštrek, M., & Marinko, M. (2018). Continuous blood pressure estimation from PPG signal. *Informatica (Slovenia)*.
- Slapnicar, G. (2018) Continuous blood pressure estimation from PPG signal. Master Thesis, Slovenia
- So, C. F., Choi, K. S., Wong, T. K. S., & Chung, J. W. Y. (2012). Recent advances in noninvasive glucose monitoring. In *Medical Devices: Evidence and Research*.
- Sobel, S. I., Chomentowski, P. J., Vyas, N., Andre, D., & Toledo, F. G. S. (2014). Accuracy of a novel noninvasive multisensor technology to estimate glucose in diabetic subjects during dynamic conditions. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/1932296813516182>
- Song, J., Li, D., Ma, X., Teng, G., & Wei, J. (2017). A robust dynamic heart-rate detection algorithm framework during intense physical activities using photoplethysmographic signals. *Sensors (Switzerland)*.
<https://doi.org/10.3390/s17112450>
- Souza, T. T., Balasingam, B., & Maev, R. (2019). Autonomous Heart Rate Tracking Methodology Using Kalman Filter and the em Algorithm. *FUSION 2019 - 22nd International Conference on Information Fusion*.
- Stevanoski, G., Achkoski, J., Koceski, S., Bogdanova, A. M., & Bogdanoski, M. (2015). Addressing communication security issues in BAN medical system: SIARS. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. https://doi.org/10.1007/978-3-319-27072-2_28
- Stevanoski, G., Kocev, I., Achkoski, J., Koceski, S., & Temelkovski, B. (2016).

- Implementation of a system for physiological status monitoring by using tactical military networks. *Defence Science Journal*. <https://doi.org/10.14429/dsj.66.9920>
- Stojanova, A., Koceski, S., & Koceska, N. (2019) Continuous Blood Pressure Monitoring as a Basis for Ambient Assisted Living (AAL) – Review of Methodologies and Devices *J Med Syst* 43, 24. <https://doi.org/10.1007/s10916-018-1138-8>
- Sun, J. X., Reisner, A. T., & Mark, R. G. (2006). A signal abnormality index for arterial blood pressure waveforms. *Computers in Cardiology*.
- Takahashi, M., Heo, Y. J., Kawanishi, T., Okitsu, T., & Takeuchi, S. (2013). Portable continuous glucose monitoring systems with implantable fluorescent hydrogel microfibers. *Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*. <https://doi.org/10.1109/MEMSYS.2013.6474439>
- Tamura, T., Maeda, Y., Sekine, M., & Yoshida, M. (2014). Wearable photoplethysmographic sensors—past and present. In *Electronics* . <https://doi.org/10.3390/electronics3020282>
- Taneja, M., & Davy, A. (2016). Resource Aware Placement of Data Analytics Platform in Fog Computing. *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2016.08.295>
- Tanweer, K. T., Hasan, S. R., & Kamboh, A. M. (2017). Motion artifact reduction from PPG signals during intense exercise using filtered X-LMS. *Proceedings - IEEE International Symposium on Circuits and Systems*. <https://doi.org/10.1109/ISCAS.2017.8050418>
- Teng, X. F., Poon, C. C. Y., Zhang, Y. T., & Bonato, P. (2008). Wearable Medical Systems for p-Health. *IEEE Reviews in Biomedical Engineering*. <https://doi.org/10.1109/RBME.2008.2008248>
- Tăuțan, A. M., Young, A., Wentink, E., & Wieringa, F. (2015). Characterization and reduction of motion artifacts in photoplethysmographic signals from a wrist-worn device. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. <https://doi.org/10.1109/EMBC.2015.7319795>
- Thomas, S. S., Nathan, V., Zong, C., Akinbola, E., Aroul, A. L. P., Philipose, L., Soundarapandian, K., Shi, X., & Jafari, R. (2014). BioWatch - A wrist watch based signal acquisition system for physiological signals including blood pressure. *2014 36th*

- Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. <https://doi.org/10.1109/EMBC.2014.6944076>
- Tierney, M. J., Tamada, J. A., Potts, R. O., Jovanovic, L., & Garg, S. (2001). Clinical evaluation of the GlucoWatch® biographer: A continual, non-invasive glucose monitor for patients with diabetes. *Biosensors and Bioelectronics*. [https://doi.org/10.1016/S0956-5663\(01\)00189-0](https://doi.org/10.1016/S0956-5663(01)00189-0)
- Trajkovik, V., Koceski, S., Vlahu-Gjorgievska, E., & Kulev, I. (2015). *Evaluation of Health Care System Model Based on Collaborative Algorithms*. https://doi.org/10.1007/978-3-319-12817-7_19
- Trajkovik, V., Vlahu-Gjorgievska, E., Koceski, S., & Kulev, I. (2015). General assisted living system architecture model. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. https://doi.org/10.1007/978-3-319-16292-8_24
- Turner, J. R., Viera, A. J., & Shimbo, D. (2015). Ambulatory blood pressure monitoring in clinical practice: A review. In *American Journal of Medicine*. <https://doi.org/10.1016/j.amjmed.2014.07.021>
- Übeyli, E. D. (2008). Feature extraction for analysis of EKG signals. *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08 - "Personalized Healthcare through Technology."* <https://doi.org/10.1109/iembs.2008.4649347>
- Vieira, J. A. P. (2011). Algorithm development for physiological signal analysis and cardiovascular disease diagnosis - A data mining approach. *Physics Departament of the Faculty of Sciences and Technology*.
- Vlahu-Gjorgievska, E., Koceski, S., Kulev, I., & Trajkovic, V. (2016). Connected-Health Algorithm: Development and Evaluation. *Journal of Medical Systems*. <https://doi.org/10.1007/s10916-016-0466-9>
- Wac, K., & Tsiourti, C. (2014). Ambulatory assessment of affect: Survey of sensor systems for monitoring of autonomic nervous systems activation in emotion. *IEEE Transactions on Affective Computing*. <https://doi.org/10.1109/TAFFC.2014.2332157>
- Wang, Lei, Lo, B. P., & Yang, G. Z. (2007). Multichannel reflective PPG earpiece sensor with passive motion cancellation. *IEEE Transactions on Biomedical Circuits and*

- Systems*. <https://doi.org/10.1109/TBCAS.2007.910900>
- Wang, Ludi, Zhou, W., Xing, Y., & Zhou, X. (2018). A novel neural network model for blood pressure estimation using photoplethysmography without electrocardiogram. *Journal of Healthcare Engineering*. <https://doi.org/10.1155/2018/7804243>
- Wang, M., Li, Z., Zhang, Q., & Wang, G. (2019). Removal of motion artifacts in photoplethysmograph sensors during intensive exercise for accurate heart rate calculation based on frequency estimation and notch filtering. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19153312>
- Wang, R., Jia, W., Mao, Z. H., Sclabassi, R. J., & Sun, M. (2014). Cuff-free blood pressure estimation using pulse transit time and heart rate. *International Conference on Signal Processing Proceedings, ICSP*. <https://doi.org/10.1109/ICOSP.2014.7014980>
- Wang, Y. (2015). Cloud-dew architecture. *International Journal of Cloud Computing*. <https://doi.org/10.1504/ijcc.2015.071717>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. In *Data Mining: Practical Machine Learning Tools and Techniques*. <https://doi.org/10.1016/c2009-0-19715-5>
- World Health Organization. (2013). A global brief on Hypertension - World Health Day 2013. In *World Health Organization*. <https://doi.org/10.1136/bmj.1.4815.882-a>
- Wu, C. M., Chuang, C. Y., Chen, Y. J., & Chen, S. C. (2016). A new estimate technology of non-invasive continuous blood pressure measurement based on electrocardiograph. *Advances in Mechanical Engineering*. <https://doi.org/10.1177/1687814016653689>
- Xie, Q., Wang, G., Peng, Z., & Lian, Y. (2019). Machine Learning Methods for Real-Time Blood Pressure Measurement Based on Photoplethysmography. *International Conference on Digital Signal Processing, DSP*. <https://doi.org/10.1109/ICDSP.2018.8631690>
- Xu, B., Xu, L. Da, Cai, H., Xie, C., Hu, J., & Bu, F. (2014). Ubiquitous data accessing method in iot-based information system for emergency medical services. *IEEE Transactions on Industrial Informatics*. <https://doi.org/10.1109/TII.2014.2306382>
- Xu, P. J., Zhang, H., & Tao, X. M. (2008). Textile-structured electrodes for

- electrocardiogram. *Textile Progress*. <https://doi.org/10.1080/00405160802597479>
- Ye, Y., He, W., Cheng, Y., Huang, W., & Zhang, Z. (2017). A robust random forest-based approach for heart rate monitoring using photoplethysmography signal contaminated by intense motion artifacts. *Sensors (Switzerland)*. <https://doi.org/10.3390/s17020385>
- Yi, S., Li, C., & Li, Q. (2015). A survey of fog computing: Concepts, applications and issues. *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*. <https://doi.org/10.1145/2757384.2757397>
- Yilmaz, T., Foster, R., & Hao, Y. (2010). Detecting vital signs with wearable wireless sensors. In *Sensors*. <https://doi.org/10.3390/s101210837>
- Yousefi, R., Nourani, M., Ostadabbas, S., & Panahi, I. (2014). A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors. *IEEE Journal of Biomedical and Health Informatics*. <https://doi.org/10.1109/JBHI.2013.2264358>
- Zhang, Q., Zhou, D., & Zeng, X. (2017). Highly wearable cuff-less blood pressure and heart rate monitoring with single-arm electrocardiogram and photoplethysmogram signals. *BioMedical Engineering Online*. <https://doi.org/10.1186/s12938-017-0317-z>
- Zhang, Yangsong, Liu, B., & Zhang, Z. (2015). Combining ensemble empirical mode decomposition with spectrum subtraction technique for heart rate monitoring using wrist-type photoplethysmography. *Biomedical Signal Processing and Control*. <https://doi.org/10.1016/j.bspc.2015.05.006>
- Zhang, Yifan, Song, S., Vullings, R., Biswas, D., Simões-Capela, N., Van Helleputte, N., Van Hoof, C., & Groenendaal, W. (2019). Motion artifact reduction for wrist-worn photoplethysmograph sensors based on different wavelengths. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19030673>
- Zhang, Z. (2015). Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2015.2406332>
- Zhang, Z., Pi, Z., & Liu, B. (2015). TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. *IEEE Transactions on Biomedical Engineering*. <https://doi.org/10.1109/TBME.2014.2359372>

Прилог 1

Листа на кратенки:

AAC	Area Above the Curve
AAL	Ambient Assisted living
AALS	Ambient Assisted Living System, систем за амбиентално помогнато живеење
AC	Alternating Current
ACI	Access Control Issues, Проблеми со контрола на пристап
AH	Account Hijacking, Киднапирање на сметка
AI	Augmentation Index (Индекс на зголемување)
AMPD	Automatic Multiscale Based Peak Detection
ANC	Adaptive Noise Cancellation
ANC	Active Noise Cancellation
ANN	Artificial Neural Network, вештачка невронска мрежа
ANS	Автономен нервен систем
APT	Advance Persistent Threats, Напредни постојани закани
ARFF	Attribute-Relation File Format
ASP	Application Service Provider
AT	Adaptive Tracking
a.u	arbitrary units
AUC	Area Under the Curve
BAN	Body Area Network, мрежа во подрачјето на телото
BP	Blood pressure, крвен притисок
BPM	Beats per minute, отчукувања во минута
CDN	Content delivery networks, мрежите за доставување на содржини
CPS	Cyber physical systems, сајбер физички системи
CVD	Cardiovascular disease

DAG	Directed Acyclic Graph
DB	Data Breaches, Повреда на податоците
DBP	Diastolic blood pressure, дијастолен крвен притисок
DC	Direct Current component
DC	Dew Computing
DL	Data Loss, Губење на податоци
DoS	Denial of Service, Негирање на услугата
DTR	Decision Tree Regression
EC	Edge Computing
ECG	Electrocardiogram
EEMD	ensemble Empirical Mode Decomposition
EMD	Empirical Mode Decomposition
EP	Elastomeric plethysmography, еластомерна плетизмографија
FaaS	Fog as a Service, Магла како услуга
FDC	Fog - dew computing
FFT	Fast Fourier Transform
GPUs	Graphics Processing Units
GSM	Global System for Mobile communications
GSR	Galvanic skin response, галвански одговор на кожата
HER	Electronic Health Record
HMI	Human to machine interactions, човек до машина интеракција
HR	Heart rate, брзината на срцеви отчукувања
HTTPS	Hypertext Transfer Protocol Secure
IaaS	Infrastructure as a Service, инфраструктура како услуга
ICA	Independent Component Analysis
IDD	Insufficient Due Diligence, Недоволна длабинска пристојност
IFCIoT	Integrated Fog Cloud IoT, инетгрирани Магла Облак интернет на нештата
IMF	Intrinsic Mode Function
InPs	Internet infrastructure provider
IoT	Internet of Things, интернет на нештата

IP	Impedance Plethysmography, импедансна плетизмографија
IP	Internet Protocol
JOSS	JOint Sparse Spectrum reconstruction
jRip	java RIPPER
LAN	Local Area Network, локална мрежа
LASI	Large Artery Stiffness Index
LED	Light- emitting diode
LMS	Least Mean Square
LMS	Local Maxima Scalogram
M2M	Machine to machine communication комуникација машина до машина
MA	Motion Artifacts, движечки артефакти
MAR	Motion Artifact Removal
MBP	Mean blood pressure (среден крвен притисок)
MCC	Mobile Cloud Computing
MEC	Mobile Edge Computing
MI	Malicious Insider, Злонамерен инсајдер
ML	Machine Learning
MLP	Multi Layer Perceptron
MLR	Multiple Linear Regression
MMV	Multiple Measurement Vector
MSE	Mean Square Error
NFV	Network function virtualisation, виртуелизација на мрежна функција
NN	Neuron Network, Невронски мрежи
OpenFog RA	OpenFog референтната архитектура
PaaS	Platform as a Service, платформа како услуга
PAC	Programmable automation controllers, програмабилни автоматизирани контролори
PAN	Personal Area Network, персонална мрежа
PCO ₂	Парцијален притисок на јаглерод диоксид
PD	Photodetector, Фотодетектор

PPG	Photoplethysmogram, фотоплетизмографија
PTT	Pule Transit Time, време на транзиција на бранот
PWV	Pulse wave velocity, брзината на пулсниот бран
QoS	Quality of service, квалитет на сервис
RAN	Radio access network
RAS	Reliability, Availability, Serviceability, надежност, достапност, услужливост
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest, Случајни шуми
RIP	Respiratory inductive plethysmography, респираторна индуктивна плетизмографија
RIPPER	Repeated Incremental Pruningto Produce Error Reduction
RLR	Regularized Linear Regression
RLS	Recursive least squares
RoC	Removing of Components
RR	Respiration rate, брзината на дишење
SaaS	Software as a Service, софтвер како услуга
SBP	Systolic blood pressure, систолен крвен притисок
SCALE	Security, Cognition, Agility, Latency, Efficiency, безбедност, сознание, агилност, латентност, ефикасност
SCICA	Single Channel Independent Component Analysis
SDN	Software defined networking, софтверско дефинирано мрежно работење
SLA	Service level agreement
SNR	Signal to Noise ratio
SpO2	Blood Oxygen Saturation, кислородна сатурација на крвта
SPT	Spectral Peak Tracking
SS	Spectrum Subtraction
SSA	Singular Spectrum Analysis
SSE	Sum of Squared Errors

SSM	Slope Sum Method
SSR	Sparse Signal Reconstruction
STI	Shared Technology Issues, Проблеми со делење на технологијата
SVD	Singular Value Decomposition
SVM	Single Measurement Vector
SVM	Support Vector Machines, Машини со носечки вектори
TROIKA	decomposiTion, sparse signal RecOnstruction, and spectral peak tracking
UE	User equipment, корисничка опрема
V2V	Vehicle to Vehicle communication – комуникација возило до возило
WAN	Wide Area Network, широка мрежа
WEKA	Waiakato Enviroment for Knowledge Analysis
WSN	Wireless sensor network, безжична сензорска мрежа
EKG	Електрокардиограм
ИКТ	Информациско комуникациски технологии
ИТ	Информациони технологии
СЗО	Светска здравствена организација